



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 162984

TO: Phuong Bui  
Location: REM-2A15&2C18  
Art Unit: 1638  
Friday, September 02, 2005  
Case Serial Number: 10/762049

From: Barb O'Bryen  
Location: Biotech-Chem Library  
Remsen 1a69  
Phone: 571-272-2518

barbara.obryen@uspto.gov

### Search Notes

**THIS PAGE BLANK (USPTO)**





# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 165364

TO: Phuong Bui  
Location: rem/2A15/2C18  
Art Unit: 1638

*Sept 14*, 2005

Case Serial Number: *10/762049*  
~~10/62049~~

From: P. Sheppard  
Location: Remsen Building  
Phone: (571) 272-2529

sheppard@uspto.gov

### Search Notes

*Interference*

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 23:27:41 ; Search time 10582 Seconds  
(without alignments)  
11214.026 Million cell updates/sec

Title: US-10-762-049-17  
Perfect score: 2449  
Sequence: 1 GCACGAGCTAGCTCGACAT.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hcg.\*
- 3: gb\_in.\*
- 4: gb\_cm.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2449	100.0	2449	6	AR477464
2	1042.8	42.6	2045	8	D89631
3	1041.2	42.5	2266	8	AY099873
4	1032	42.1	2193	8	BN4581745
5	1018.2	41.6	1977	6	AX506807
6	1018.2	41.6	1977	6	AX506807
7	897.4	36.6	2037	8	AB004060
8	864.8	35.3	2219	8	AJ601439
9	846	34.5	1974	6	AX651455
10	846	34.5	2327	8	AK066932
11	765.2	31.2	1981	6	AR477457
12	720.6	29.4	2067	6	AR477463
13	681	27.8	2259	8	AJ704374
14	678.6	27.7	2256	8	AB054645
15	677.4	27.7	2640	8	AF355602
16	667.8	27.3	2181	8	AK104831
17	666.2	27.2	2338	8	AK072809
18	660.2	27.0	1896	6	AX651700
19	659.2	26.9	1917	8	AB023423

20	659.2	26.9	2470	8	AF347614	AF347614	Lycopersi
21	655.4	26.8	2279	6	AR477456	AR477456	Sequence
22	653.4	26.7	2159	8	AB042322	AB042322	Arabidops
23	653.4	26.7	2439	8	AF309643	AF309643	Solanum t
24	650.4	26.6	1989	6	AX653919	AX653919	Sequence
25	650.4	26.6	2791	8	AK103007	AK103007	Oryza sat
26	649.8	26.5	2262	8	BN4416460	AJ416460	Brassica
27	647.2	26.4	1947	6	AX653223	AX653223	Sequence
28	647.2	26.4	1989	8	AF493792	AF493792	Oryza sat
29	646	26.4	2167	8	BN4311388	AJ311388	Brassica
30	640.4	26.1	2310	8	BT009249	BT009249	Triticum
31	640.4	26.1	2311	6	AR477465	AR477465	Sequence
32	639	26.1	2248	8	AF332959	AF332959	Lycopersi
33	637.4	26.0	2327	8	AF347613	AF347613	Lycopersi
34	634.4	25.9	1971	6	BD175021	BD175021	Gene enco
35	634.4	25.9	1971	8	AB049624	AB049624	Arabidops
36	632.6	25.8	2494	8	AK067270	AK067270	Oryza sat
37	630.4	25.7	1940	6	AX654768	AX654768	Sequence
38	627	25.6	2442	8	HVHVS11	X96431	Hordum vul
39	622	25.4	2337	8	SHSHST2	X82256	S.hamata SH
40	620.6	25.3	2456	8	HVU52867	U52867	Hordeum vul
41	614	25.1	2454	8	SHSHST1	X82255	S.hamata SH
42	610	24.9	2515	8	SSSULPTRP	X96761	S.stapfianu
43	609.8	24.9	2313	8	AB018695	AB018695	Arabidops
44	600.2	24.5	2000	6	AX653563	AX653563	Sequence
45	586.4	23.9	2808	8	AK121195	AK121195	Oryza sat

ALIGNMENTS

RESULT 1  
AR477464  
LOCUS AR477464 2449 bp DNA linear PAT 14-MAY-2004  
DEFINITION Sequence 17 from patent US 6696292.  
ACCESSION AR477464  
VERSION AR477464.1 GI:47234922  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2449)  
AUTHORS Allen,S.M., Falco,S.C. and Thorpe,C.J.  
TITLE Genes encoding sulfate assimilation proteins  
JOURNAL Patent: US 6696292-A 17 24-FEB-2004;  
FEATURES  
Location/Qualifiers  
source 1..2449  
/organism="unknown"  
/mol\_cype="genomic DNA"

Query Match	100.0%	Score 2449;	DB 6;	Length 2449;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2449;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GCACGAGCTAGCTCGCACATTAAGTTATATACATATTTGCTTCTTAGAAATACTAT	60	
Db	1	GCACGAGCTAGCTCGCACATTAAGTTATATATACATATTTGCTTCTTAGAAATACTAT	60	
Qy	61	TATTGAAGATATGGGGAGTGATATTCAGTACCTTTGGGCATGAACTTTCAGAG	120	
Db	61	TATTGAAGATATGGGGAGTGATATTCAGTACCTTTGGGCATGAACTTTCAGAG	120	
Qy	121	AGTGACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAGAGTCTCTAAAGTACTCTTT	180	
Db	121	AGTGACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAGAGTCTCTAAAGTACTCTTT	180	
Qy	181	GAAGGAGACTTCTTCCCTGATGACCCCTTTGAGGAGTTCAGAAACAGCCAGCTTCCAA	240	
Db	181	GAAGGAGACTTCTTCCCTGATGACCCCTTTGAGGAGTTCAGAAACAGCCAGCTTCCAA	240	
Qy	241	GAAGTTCAGCTTGGCCTTCAGTCTTCTTCCCATTTTCGAATGGCTCCCAATACAC	300	

Dbb 241 GAAGTTCATGCTTGGCGCTTCAGTTCCTCTTCCCAATTTTCGAATGGCTCCCAAAATACAC 300  
Qy 301 CTTTCAGTTCCTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTC 360  
Db 301 CTTTCAGTTCCTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTC 360  
Qy 361 TCAGGGATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATTCCTTGGACTATATTCGAG 420  
Db 361 TCAGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATTCCTTGGACTATATTCGAG 420  
Qy 421 CTTTATACCAACCATTCATTTATCGGATGATGGGTAGCTCGAGGATTTGGCAGTGGGAC 480  
Db 421 CTTTATACCAACCATTCATTTATCGGATGATGGGTAGCTCGAGGATTTGGCAGTGGGAC 480  
Qy 481 TGTGGCGGTTGGATCGCTTCGTATGGGTTTCGATGTTGAGTAATGCCGTTGATCCCAATGA 540  
Db 481 TGTGGCGGTTGGATCGCTTCGTATGGGTTTCGATGTTGAGTAATGCCGTTGATCCCAATGA 540  
Qy 541 AGACCCAAAGCTTTA CTTCCACTGGCTTTCACAGCTACATTTATTCGTGGTGTTC 600  
Db 541 AGACCCAAAGCTTTA CTTCCACTGGCTTTCACAGCTACATTTATTCGTGGTGTTC 600  
Qy 601 GGCCTGCTGGGTCCTGTTAGGTTGGGTTGATCGTGGATTTCTGTCACATGCCAAT 660  
Db 601 GGCCTGCTGGGTCCTGTTAGGTTGGGTTGATCGTGGATTTCTGTCACATGCCAAT 660  
Qy 661 AATAGGGTTTCATGGGAGGAGCAGCCACGGTGGTGTCTGCAGCAACTAAATTCGATTC 720  
Db 661 AATAGGGTTTCATGGGAGGAGCAGCCACGGTGGTGTCTGCAGCAACTAAATTCGATTC 720  
Qy 721 TGGCCTTGAGCATTTCCACCATGGAGCTGATATCATATCAGTGAATGGCTCTGTTTTCAC 780  
Db 721 TGGCCTTGAGCATTTCCACCATGGAGCTGATATCATATCAGTGAATGGCTCTGTTTTCAC 780  
Qy 781 CCAAACTCATGAGTGGAGTGGAAAGTCTGTGTTAGGATGCTTTCATTTCTTCCT 840  
Db 781 CCAAACTCATGAGTGGAGTGGAAAGTCTGTGTTAGGATGCTTTCATTTCTTCCT 840  
Qy 841 CCTTAGCACAAGATACTTCAGCAAAAAACGACCAAGGTTTTTGGGTGTGAGCAATGGC 900  
Db 841 CCTTAGCACAAGATACTTCAGCAAAAAACGACCAAGGTTTTTGGGTGTGAGCAATGGC 900  
Qy 901 GCCATTGACGTCGGTTATATTTGGAAAGTCTCTTGGTTTATTTCACTCA CGCGAGAAGCA 960  
Db 901 GCCATTGACGTCGGTTATATTTGGAAAGTCTCTTGGTTTATTTCACTCA CGCGAGAAGCA 960  
Qy 961 CGGTGTTGAAGTGATAGGAGAACTGAAGAGGGTTTGAATCCACCCTCACTCAAAATCT 1020  
Db 961 CGGTGTTGAAGTGATAGGAGAACTGAAGAGGGTTTGAATCCACCCTCACTCAAAATCT 1020  
Qy 1021 GGTATTGTTGTCGCCCTTACATGACTACAGCTGTCAAAACCTGGCATTTGCTGGCATCAT 1080  
Db 1021 GGTATTGTTGTCGCCCTTACATGACTACAGCTGTCAAAACCTGGCATTTGCTGGCATCAT 1080  
Qy 1081 ATCATTGCGGAAAGGAATAGCAGTAGGAAGAGCTTTTGGCAATGTATAAAAAATTA 1140  
Db 1081 ATCATTGCGGAAAGGAATAGCAGTAGGAAGAGCTTTTGGCAATGTATAAAAAATTA 1140  
Qy 1141 TGATGCCAACAAGAGATCATAGCTATTCGGGACCATGAACGTAGTTGGTCTTTCACCTC 1200  
Db 1141 TGATGCCAACAAGAGATCATAGCTATTCGGGACCATGAACGTAGTTGGTCTTTCACCTC 1200  
Qy 1201 TTGCTACCTCACAAAGGACCAATTTTCGGGTTCCGGCTGTGAACATAAACGCTGGATGCAA 1260  
Db 1201 TTGCTACCTCACAAAGGACCAATTTTCGGGTTCCGGCTGTGAACATAAACGCTGGATGCAA 1260  
Qy 1261 GACAGAGCTTCCAAATTAATGTCATTCGAGTAATGTTGACATTTGTTATTCCTGAC 1320  
Db 1261 GACAGAGCTTCCAAATTAATGTCATTCGAGTAATGTTGACATTTGTTATTCCTGAC 1320  
Qy 1321 ACCCTGTTTCCATTACACTCCCTGGTGGTGTATCAGCTATTATCGTATCTGCAATGCT 1380  
Db 1321 ACCCTGTTTCCATTACACTCCCTGGTGGTGTATCAGCTATTATTCGTATCTGCAATGCT 1380

Qy 1381 TGGACTCATAGATTATGAAGCAGCCCATCCATCTATTTAAGGTTGACAAAATTTGACTTTGT 1440  
Db 1381 TGGACTCATAGATTATGAAGCAGCCCATCCATCTATTTAAGGTTGACAAAATTTGACTTTGT 1440  
Qy 1441 GGTGTGCATGAGTGATGATACATTTGGGCTGCTCTTTGGCAGTGTGAAATTTGGCTTAGTCAT 1500  
Db 1441 GGTGTGCATGAGTGATGATACATTTGGGCTGCTCTTTGGCAGTGTGAAATTTGGCTTAGTCAT 1500  
Qy 1501 AGCTATGTGTAATATCTGTACTTTCGGGTACTTCTATTTATTTGCAAGGCCAAGGACATTCGT 1560  
Db 1501 AGCTATGTGTAATATCTGTACTTTCGGGTACTTCTATTTATTTGCAAGGCCAAGGACATTCGT 1560  
Qy 1561 TTTTGGGCAACATTTCCAAATTTCTGTATATACCGAAATGTTGAGCACTATCAAAATCCAAA 1620  
Db 1561 TTTTGGGCAACATTTCCAAATTTCTGTATATACCGAAATGTTGAGCACTATCAAAATCCAAA 1620  
Qy 1621 ACATGTTCTGGAATGCTTAATTTCTAGAGATTCAGATGATGACCAATTTTAAAGCTAC 1680  
Db 1621 ACATGTTCTGGAATGCTTAATTTCTAGAGATTCAGATGATGACCAATTTTAAAGCTAC 1680  
Qy 1681 CTATTTAAGAGAAAGGATCACAAGGTGGATTCATGAAGAAAGAAAGAAATTTAAAGCTAC 1740  
Db 1681 CTATTTAAGAGAAAGGATCACAAGGTGGATTCATGAAGAAAGAAAGAAATTTAAAGCTAC 1740  
Qy 1741 AGGGGAGACTAGTTTGCAGTATGTTATATTCATATGATGCTGTTTGGAAAACATTTGATAC 1800  
Db 1741 AGGGGAGACTAGTTTGCAGTATGTTATATTCATATGATGCTGTTTGGAAAACATTTGATAC 1800  
Qy 1801 AAGTGGAAATAGATGCTTTGAAAGAGGTGAAGAAATTAAGAGAGAGAGAGTACAGCT 1860  
Db 1801 AAGTGGAAATAGATGCTTTGAAAGAGGTGAAGAAATTAAGAGAGAGAGAGTACAGCT 1860  
Qy 1861 TGTGTTTGGTCAATCTGTAAGTGAAGTGAAGAACTGGAACAACTCGAAGTTTCCAAA 1920  
Db 1861 TGTGTTTGGTCAATCTGTAAGTGAAGTGAAGAACTGGAACAACTCGAAGTTTCCAAA 1920  
Qy 1921 TCATTTAGGGAAGAAATGGATCTATCTGACTGTTGAAGAGGCGGTTGGAGCATGCAACTT 1980  
Db 1921 TCATTTAGGGAAGAAATGGATCTATCTGACTGTTGAAGAGGCGGTTGGAGCATGCAACTT 1980  
Qy 1981 CAATCTAGTGCAGCAAAAAACGAAACCCAAAGAAAGTGAAGAGAGGTTGGAACAATGT 2040  
Db 1981 CAATCTAGTGCAGCAAAAAACGAAACCCAAAGAAAGTGAAGAGAGGTTGGAACAATGT 2040  
Qy 2041 GTGACTGAGTCATATGCCAAAGAGTATCTAAATAACTCAAAAAGCTTATTCGTTTCT 2100  
Db 2041 GTGACTGAGTCATATGCCAAAGAGTATCTAAATAACTCAAAAAGCTTATTCGTTTCT 2100  
Qy 2101 CTTAGTAATGTTTACCCTACATATGTTGGCATGAGAAATTTCTGAATCAGCCGAAGAAGT 2160  
Db 2101 CTTAGTAATGTTTACCCTACATATGTTGGCATGAGAAATTTCTGAATCAGCCGAAGAAGT 2160  
Qy 2161 TTTAAAGCATAGGAAATGAAAGATGCAGAGGCTCTTAATTTCTCAACTCTGCATCT 2220  
Db 2161 TTTAAAGCATAGGAAATGAAAGATGCAGAGGCTCTCTTAATTTCTCAACTCTGCATCT 2220  
Qy 2221 TAGTTAGAGAAATCTCCTATGATAGCTGTTGAAATTAATCTTTAGTATCATGCTTG 2280  
Db 2221 TAGTTAGAGAAATCTCCTATGATAGCTGTTGAAATTAATCTTTAGTATCATGCTTG 2280  
Qy 2281 ATAATATATTTCAAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACAATG 2340  
Db 2281 ATAATATATTTCAAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACAATG 2340  
Qy 2341 TAAAGAGTAAAGAGTGTGTTAGCACTCTCCATCATTAATTTGAAAGTAAATTTGCC 2400  
Db 2341 TAAAGAGTAAAGAGTGTGTTAGCACTCTCCATCATTAATTTGAAAGTAAATTTGCC 2400  
Qy 2401 ATGAGAAATTTAAATCTTTTGGAAAAAATTTTGGAAAAAATTTTGGAAAAAATTTTGG 2449  
Db 2401 ATGAGAAATTTAAATCTTTTGGAAAAAATTTTGGAAAAAATTTTGGAAAAAATTTTGG 2449

RESULT 2	D89631	2045 bp	mrna	linear	PLN 14-FEB-2004
LOCUS	Arabidopsis thaliana mRNA for sulfate transporter, complete cds.				
DEFINITION	D89631				
ACCESSION	D89631				
VERSION	D89631.1	GI:2285884			
KEYWORDS	ATST1; sulfate transporter.				
SOURCE	Arabidopsis thaliana				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1	Yamaguchi, Y., Nakamura, T., Harada, E., Koizumi, N. and Sano, H.			
AUTHORS	The electronic Plant Gene Register				
TITLE	Plant Physiol. 113 (4), 1463-1465 (1997)				
JOURNAL	97267155				
MEDLINE	2	(bases 1 to 2045)			
REFERENCE	Yamaguchi, Y.				
AUTHORS	Direct Submission				
TITLE	Submitted (29-NOV-1996) Yube Yamaguchi, Nara Institute of Science and Technology; 8916-5 Takayama, Ikoma, Nara 630-01, Japan				
JOURNAL	(E-mail: yu-yama@bsmailgate.aist-nara.ac.jp, Tel: 07437-2-5652, Fax: 07437-2-5659)				
FEATURES	Location/Qualifiers				
source	1. .2045				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				
	/db_xref="taxon:3702"				
	/tissue_type="above-ground parts"				
	/clone_lib="lambda ZAP II"				
	/ecotype="landsberg erecta"				
gene	1. .2045				
	/gene="ATST1"				
CDS	33. .2009				
	/genes="ATST1"				
	/codon_start=1				
	/product="sulfate transporter"				
	/protein_id="BA21657.1"				
	/db_xref="GI:2285885"				
	/translations="MGTEDTYFQGAEEHRRHHTVEAPQPQFLSKLOYSVKETLFPD DDPFPQKQNAQSKFLVGLFLPIPEWAPRYNLKPFKSDLIAGITIASLAIPOGIS YAKLNLPTILGYSFSPPLVAVLGGSSDLAVGTAVASLLTGAMLSKEVDAEKDP KLYHLAFTATFAGVLEASLIGLFRGLFVDFLSHATVFGMGAATVSLQIKGIF MLPFTSDTIVSLVMSRVSGQHEWRNESVGLCGFLFLLSITRYFSIKPKPFVWAA GAKTDSGLSLVYTHAERHQVQVIGDLKGLNPLSGDLITSPYKSTAVTKGLI TGIALAAGIAGVRSFMKYNINDGNKEMIAFCMMNIVGSFTSCVLTGTFPFSRAVN YNAGCTAMSNIVMAIATMFTLLPLTFLPHTPLVLSLIIISAMGLGIDYQAAILHL KVDKDFLCVMSAVGVVFGSVRIGLVVAISIALRLFPVSPKTAVKGNIPNSMIY RNTQVPSSTVTFGLLIDDAIDYFANASYLRRRIIRWIDDEBERVKQSGESLQVI ILMDSVGNIDTSGILSNMVEIKKVIDRALKVLVSNPKGBEIVKKLIRSKFPIGHLGKRE WMFLTVEAVEARCSYMLHTFKTEPPASKNEPNNV"				
ORIGIN					
Query Match	42.6%	Score 1042.8;	DB 8;	Length 2045;	
Best Local Similarity	72.4%	Pred. No. 3.9e-248;			
Matches 1366;	Conservative 0;	Mismatches 517;	Indels 3;	Gaps 1;	
Qy	125	CACCAAGTCGAGGTTCCACCGCCACACCGCTTTTTCAAAGTCTCTAAAGTACTCTTTTCAAG 184			
Db	90	CACACGGTGAGGCTCCACAACCTCAACCGTCTCTGAAGTCACCTTCAGTACTCAGTGAAG 149			
Qy	185	GAGACTTTCCTCGATGACCTTTTGAGGAGAGTTCAAGAACACGCGCTTCCAGAAG 244			
Db	150	GAACCTCTGTTTCCAGACGACCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAA 209			
Qy	245	TTCAATGCTGCSCCTTCAGTTCTCTCCCACTTTTCGAATGGGCTCCCAAAATACACCTTT 304			
Db	210	TTTGTGTAGGCTTCAATACTTCTCCCGAATTTTCGAATGGGCACACCGCTACAAATCTC 269			
Qy	305	CAGTTCTTGAAGAGTGACCTCATAGCTGGCATCACCAATCGTAGCTTGGCCATTCCTTCAG 364			

Db	270	AAGTTCTTCAAAATCAGATCTCATCGCCGGGAATCACCAATCGCTAGTCTCGCCATCCCTCAG	329
Qy	365	GGCATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTGGACTATATTCGAGCTTT	424
Db	330	GGCATCAGTTACGCCAAACCTGTAACTTGCGCCCAATCTTGGSCCTTTATTCGAGTTTT	389
Qy	425	ATACCACCATGATTTATATGGGATGATGGGTAGCTCGAGGGATTTGGCAGCTGGGACHTGT	484
Db	390	GTACCGCCATTGGTATACGCGGTGCTAGGAGTTCAGGGACTTTAGCGGTGGGAACGGTT	449
Qy	485	GCCTTGGATCGCTCTTGATGGGTTCGATGTCGATTAATCCGTTGATCCCAATGAAGAC	544
Db	450	GCCTTGGCTCTCTGTTGACAGTGCATGCTGAGCAAGAAGTTGATGCTGAGAAAGAT	509
Qy	545	CCAAAGCTTTACCTCCACCTGGCTTTCCACAGCTACATTTATTTGCTGGTGTTTTTTCAGGCT	604
Db	510	CCTAAGCTTTACCTTCACCTTGTCTTTCACCGCCACTTTTTTCGCCGCGTTCGGAAGCC	569
Qy	605	GCCTTGGTCTGTTTAGTTGGGTGATCGTGGATTTCTGTGCATGTCGAACCATTAATA	664
Db	570	TCTCTTGGAAATTTTCAGGTTAGGGTTCATAGTGGACTTTCTATCGATGCAACGATAGTA	629
Qy	665	GGTTCATGGGAGGAGCCACGCTGGTGTGTCTGCAGCAATTAATAATCGATTCCTGGC	724
Db	630	GSATTCATGGGAGGAGCAGCGGTGGTAGTCTGCACAGCTTAAGGGTATTTTCGGGA	689
Qy	725	CTTGAGCATTTACCAATGGAGCTGATATCATATCAGTGTGCTGCAGCAATTAATAATCGATTCCTGGC	784
Db	690	CTTAAACATTTACAGACTCTACCGATGTTATCTCTGTCTATGGTTCGCTTTTCTCCCA	749
Qy	785	ACTCATGAGTGGAGGTGGAAAGTGTGTGTAGGATGTCTTCAATTTTCTTCTCTCTT	844
Db	750	ACTCACGAGTGGAGATGGGAAAGTGGCGTCTCTTTGGATGTGTTTCTATTCTTCTTCTC	809
Qy	845	AGCACAGACTTTCAGCAAAAAACGACCAAGGTTTTTTTTGGGTGTCAGCAATGGCGCCA	904
Db	810	TCACACAGATATTTTCAGCATCAAGAAACCAAAAAATCTTTTTGGGTGGCGCGATGGCTCT	869
Qy	905	TTGACGTCCGTATATTTGGGAAGTCTCTTGTTTTATTTCATCAACGCGGAGAGACAGGT	964
Db	870	TTGACCTCAGTGATCTTGGAGTCTCTTGTTTTATCTTCACTCAACGCGGAGAGACATGGT	929
Qy	965	GTTGAAGTGATAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAAAATCTGGTA	1024
Db	930	GTTCAAGTGATAGGGACCTGAAGAAGGTTTGAATCCACTCTCCGGTTCTGATCTCATC	989
Qy	1025	TTTGTGTCGCTTACATGACTACAGCTGTCAAAATCTGGCATTTGTTGGCATCATATCA	1084
Db	990	TTTTACTTCCCTTTACATGTCCACAGCTGTCAAAACTGGCCTCATCACTGGCATCATTTGT	1049
Qy	1085	CTTGGGGAAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATAAAATATACAATATTGAT	1144
Db	1050	CTCGCTGAAGGAATAGCAGTGGGAGAGTTTGGCATTTCAAGAACTTAAACATAGAC	1109
Qy	1145	GGCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTTTCACTCTTGC	1204
Db	1110	GGGAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTGTGTTCCATCATCTGT	1169
Qy	1205	TACCTCAACACGAGCAATTTTCGGCTTGGCTGTGAACTATTAACGGTGGATGGAAGACA	1264
Db	1170	TACCTCACACCGGACCAATTTTCAAGTCTGGCAGTGAACCTAACACGCGGGTTGCAAGAC	1229
Qy	1265	GCAGCTTCCAACATTAATGTCACTTCGAGTAATGTGTGACATTTTATTCCTGACACCC	1324
Db	1230	GCAATGTCCAACATAGTATGGGATTTGGGTATGTTTCACTCTCTTCTCTCACACCG	1289
Qy	1325	TTGCTTCATTTACACTCCCTCGTGGTGCTATCAGCTATTATCGTATCTGCAATGCTTGA	1384
Db	1290	CTTTTTCACTACACCACTCGTCTGCTCTCTCGCCATCATCATATCCGCAATGCTCGGA	1349
Qy	1385	CTCATAGTTATGAAGCAGCATCACTATTTAAGGTGGAACAAATTTGACTTTGTGGTG	1444
Db	1350	CTCATTTGACTATCAAGCTGCACTCACTCTGGAAGTTGACAAAGTTGCACTTCTCTCGTC	1409







3' UTR  
2057..-2180  
/gene="bstr3.1"

ORIGIN

Query Match 42.1%; Score 1032; DB 8; Length 2193;  
Best Local Similarity 72.2%; Pred. No. 1.9e-245;  
Matches 1357; Conservative 0; Mismatches 520; Indels 3; Gaps 1;

KLXLHLAFTATPFAGVLEASLGIFRLGFIWDFLUSHATIVGFMGGAATVVSLOOLKGF  
GLKHFTEADVISVMRSVFSQTHQWVESGVLGCCFLFLLSSTRYSPYKPKPFMWAA  
MAPLTSVLGLLVYFTHAERHGVQVIGDLKKLPLSVSLVFTSPYKSTALKCLLI  
TGIITLAEGTAVGRSFAPEKNVINDGNKSMIAFGMMNIVGSPTSCLYLTGPPRSVAV  
FNAGCKTAVSNYMAIAVMPTLLFHPFHYTFLVLLSSIIWVAMGLGLIDYQAIHLW  
KVDKDFVCMSSAYFVGSVEIGILVAVVISIARLLLFVSDPRVAVKGNIPNTMIY  
RNTDQPYSRIVFGLJLLEIDAPIYFANAGYLREIRTRWVDEEDRIKASGNSLOVY  
LIDMSAVGNIDTSGISMBEIKKIMDRRELKVLANPKGEVVKLLTRSKFPIGNLGKE  
WMFLTWGEAVEACSFMLHTSKTEPASKERPNNV"  
2057..-2180  
/gene="bstr3.1"

131 GTGAGGTTCCACGGCAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTGAAGGAGACT 190  
143 GTGGAGGCTCCGGATCCTCAGCGGTTCTTGAAGTCACTTCAGTATTCAAGTGAAGAACT 202

191 TTCTTCCCTGATCACCCTTGGGCGAGTTCAGAAACAAGCCGCTTCCAGAGTTCATG 250  
203 CTGTTTCCAGACGATCCTTTACAGAGTTCAGAAACAAGGACATCAAGACAAGTTGTA 262

251 CTGGGCTTCCAGTTCCTTCCCATTTTCGAATGGGCTCCCAAAATACACCTTTCAAGTTC 310  
263 CTAGGCTCAAAATACCTTTGCAATATGGAAATGGGCTCCACTCTACAAATTCAGACTC 322

311 TTGAAGCTGACCTCATAGCTGCGATCAACATCGCTAGCTTGGCCATTTCTCAGGCGATC 370  
323 TTTAAATACAGATCCTATCGCTGGAATCACATCGCTAGCTCGCCATCCTCAGGGCATC 382

371 AGTTATGCCAGCTCGCCACCTCCCTCCAAATCTTGGACTATATTGGAGCTTTATACCA 430  
383 AGTTACGCTAAATCGCTAACTTGGCTCCCATCTTGGGCTTTTACTCGAGTTTGTTC 442

431 CCAATTGATTTATCGGATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGACTGTGGCGGTT 490  
443 CCATTGGTGTTCGCTGTGCTGGGAGTTCAAGGAAATTTGGCGTGGGACGCTGGCGGTT 502

491 GGATCGCTTCGATAGGGTTCGATGTTGAGTAATGCGCTGATGCCAAATGAAGACCCAAAG 550  
503 GGATCTTTGTTGACAGGTGCGATGCTGAGCAAGAAGTTGACGCTGAGAAGAATCTTAAG 562

551 CTTTACTCCACCTGCTTTCACAGCTACATATTGCTGGTGTGTTTTTCAGGCTGCTTG 610  
563 CTCTACTTCATCTGCTTTTCACTGCCAGTTTTTCGCGAGGCGTTCTCGAAGCCTCTCTC 622

611 GGTCTGTTTAGTTGGGGTTGATCGTGGATTTTTCTGTCAATGCCAACCAATAATAGGGTTC 670  
623 GGCAATTTTCAGGTTAGGGTTTCAAGTGGATTTTCTATCGCATGCAACGATAGTGGGATTC 682

671 ATGGAGGAGACGACCGGCTGCTGTCAGCAACTAAATCGAATCTTTGGCCCTTGAG 730  
683 ATGGAGGAGACGACGCGGTGCTGAGTTCGCAACAGCTTAAAGGATTTTTTGGACTTAA 742

731 CATTTCCCATGAGCTGATATCATATCAGTATCGGCTCTGTTTTCACCCAACTCAT 790  
743 CATTTCCAGAAAGCCATGACGTTATTCTGTGTCATGCTTCAGTCTTCTCTCAAACTCAT 802

791 GAGTGGAGGTGGGAAAGTGTGTTGTTAGGATGTGCTTCATTTTCTTCTCTCTTAGCACA 850  
803 CAGTGGAGATGGGAGATGGTGTCTCGGCTGTGTTTCTTTCTTCTCTCTTCTTCCACC 862

851 AGATCTCTAGCAAAAAACGACCAAGGTTTTTTTGGGTTCAGCAATGGGCGCATGACG 910  
863 AGATATTTTCAGCAACGAAGAAACCAAAATTTCTTTGGGTAGTTCGATGGCTCCTTTGACC 922

911 TCCGTTTATATGGGAAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGACACGCTGTTGAA 970  
923 TCAGTGATTTTGGAAAGCTTCTTGGTGTACTTCACTCAGCGGAGAGACATGTTGTTC 982

971 GTGATAGGAGAACTGAAGAAGGGTTTGAATCCACCATCATCTCACAAATCTGTTATTTG 1030  
983 GTGATAGGGGACCTGAAGAAAGGGTTGAACCGCTCTCTGTGTGATCTTGTCTTCACT 1042

1031 TGCCTTACATGACTACAGCTGTCAAACTGGGATTTGCTGGCATCATATCATCTTGGCG 1090  
1043 TGCCTTACATGCTCAACAGCTCTCAAACTGGCCTCATCACTGGGATCATTAATCTCTT 1102

1091 GAAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATAAAAATTTACAATATTCATGSCAAC 1150  
1103 GAAGGATAGCTGTGGGAAGGAGCTTTGCCATGTTTCAAGAACTACAAATAGACGTAAC 1162

1151 AAAGGATCATAGCTATTGGGACCATGAACGCTAGTTGGTTCTTTCACTCTTTGCTACCTC 1210  
1163 AAAGAAATGATAGCGTTTGGAAATGATGAACATCGTTGGTTTCCCTTCACTCTTTGCTACCTC 1222

1211 ACAACAGGACCAATTTTGGGCTTGGCTGTGAACATTAACGCTGGATGCAAGACAGAGCT 1270  
1223 ACAACTGGAACCGTTTTCAGATCGGCGGTGAACCTTCAATCGCGGTTGCAAGACTGCCGTG 1282

1271 TCCAAATTTAATGTCACCTTCAGTAAATGTTGACATTTGTTATTTCTCTGACACCTTTGTT 1330  
1283 TCCAAATAGTAAATGGCCATTTGGGTTATGTTTCAATGCTCTTCCACAGCCTTTTCTT 1342

1331 CATTACATCCCTCGTGGTGTCTATCAGCTATTATCGTATCTGCAATGCTTGGACTCAT 1390  
1343 CACTACAGCCCTCGTGGTGTCTCTCTCCATCATCATGCTCGCAATGCTCGGACTCATC 1402

1391 GATTATGAAGACCAATCCATCTATTTAAGGTTGACAAATTTGACTTTTGGTGTGTCATG 1450  
1403 GACTATCAAGCTGCCATCCATCTTTTGAAGGTTGACAAATTCGATTTCTTCTGTCATG 1462

1451 AGTGACATCATTTGGGCTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCACTAGCTATTGTA 1510  
1463 ACGGCTCTTTTGGGCTGGTGTGGTAGTGTGCGAGATCGGACTTTGTTGCCAGTGGTA 1522

1511 ATATCTGTAATTTCCGGGTACTCTTATTTATGCAAGCCCAAGGACATTCGTTTTTGGGCAAC 1570  
1523 ATATCGATAGCAAGGTTGTGCTGTTCGTGTCGAGGCGGAGGACTCGCGTGAAGGAAAC 1582

1571 ATTCCAAATTTCTGATATATCCGAAATGTTGAGCATCAAAATGCAAAACATGTTCTCT 1630  
1583 ATACCAAAACCATGATCTATAGGAACACTGACCAAGTATCTTACTCAAGAAATTTGTTCT 1642

1631 GGAATGCTAAATCTAGAGATTGATGACCAATTTTACTTTGCAATGCGGAGCTATTTAAGA 1690  
1643 GGTCTTCTCATCTCGAGATTGATGCTCCATCTACTTTGCCAACGCTGGTTACTTGGCG 1702

1691 GAAAGGATCAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACAGGGAGACT 1750  
1703 GAGAGAAATCAAGGTGGGTGCGATGAAGAGGAAGATAGGATCAAGAGCATCAGGAGAAAT 1762

1751 AGTTTGCAGTATGTTATTAATGATATGATGCTGTTGGAAACATTTGATACAGTGGGAATA 1810  
1763 AGTTTACAATATGTTATATCTCGATATGTTGCTGTGTTGTAATATCGACACGAGTGGTATA 1822

1811 AGTATGCTTGAAGAGGTGAAGAAGATTAACAGAGAAAGAGAGCTACAGCTTCTTTTGGTTC 1870  
1823 ACATATGAGGAGAAATTAAGAAATATCATGACAGAGAGAGTTAAAGTTAGTATTGGCA 1882

1871 AATCTCTGAAGTGAAGTGAAGAAACTGAACAAATCGAAGT---TCCAAATCATTTTA 1927  
1883 AATCCAAAAGGAGAGGTGCTGGAAGAAACTAAACAGATCCAAATTCATCGGTGAAATTTG 1942

1928 GGAAGAAATGATCTATCTGATCTGTTGAAGGCGGTTGGAGCATGCACTTCAATCTA 1987  
1943 GGCNAAGAGTGGATGTTCTTAAACAGTTGGGAAGCGGTGGAAGCTTGTAGTTTCACTT 2002

1988 CGTGCAGACCAAAACGAAACC 2007  
2003 CACACGTCAAGACCGAACC 2022



RESULT 5	AX506807	AX506807	1977 bp	DNA	linear	PAT 27-SEP-2002
LOCUS	Sequence 1502 from Patent WO0216655.					
DEFINITION	AX506807					
ACCESSION	AX506807.1	GI:23388044				
VERSION						
KEYWORDS	Arabidopsis thaliana (thale cress)					
SOURCE	Arabidopsis thaliana					
ORGANISM	Arabidopsis thaliana					
REFERENCE	1					
AUTHORS	Harper, J.P., Kreps, J., Wang, X. and Zhu, T.					
TITLE	Stress-regulated genes of plants, transgenic plants containing same, and methods of use					
JOURNAL	Patent: WO 0216655-A 1502 28-FEB-2002; The Scripps Research Institute (US) ; Syngenta Participations AG (CH)					
FEATURES	Location/Qualifiers					
source	1. .1977					
	/organism="Arabidopsis thaliana"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:3702"					
ORIGIN						
Query Match	41.6%;	Score 1018.2;	DB 6;	Length 1977;		
Best Local Similarity	72.3%;	Pred. No. 5.1e-242;				
Matches 1364;	Conservative 0;	Mismatches 518;	Indels 5;	Gaps 3;		
QY	125	CACCAAGTCAGGTTCCACCGCCACACAGCCGCTTTTTCAGTCTCTAAGTACTCTTTGAG	184			
DB	58	CACACGGTGGAGGCTCCACAACTCCACACCGCTTCTTGAAGTCACTTCAGTACTCAGTGAAG	117			
QY	185	GAGACTTTCTTCCCTGATGACCTTTTGAGCAGTTTCAAGAACAAAGCCAGCTTCCAAAGAG	244			
DB	118	GAACCTCTGTTTCAGACGACCTTTTGAACAATTTAAGAACCAAAATGATCAAGAAA	177			
QY	245	TTCAATGCTGGCCCTTCAGTCTTCTTCCCAATTTTCCAAATGGGCTCCCAATACACCTTT	304			
DB	178	TTTGTGTTAGGCTCAAAATCTTCTCCGATTTTCGAATGGGCACCAGCTACAATCTC	237			
QY	305	CAGTTCTTGAAGTGACCTCATAGCTGGCATCACCATCGCTAGCTGGCCATCTCTCAG	364			
DB	238	AAGTTCTTCAAAATCAGATCTCATGCGCGGAATCACCATCGCTAGCTCGCCATCCCTCAG	297			
QY	365	GGCATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTGAGACTATATTCGAGCTTT	424			
DB	298	GGCATCAGTTACGCCAAACTTGTAACCTTGCCCCCAATCTTGCCCTTATTCGAGTTT	357			
QY	425	ATACCAACCATTTATATCGCATGATGGGTAGCTCGAGGATTTGGCAGTGGGACTGTG	484			
DB	358	GTACCGCCATTGGTATACGCGGTCTAGGAGTTCAAGGACTTAGCGGTGGGAACGGTT	417			
QY	485	CGGTTGGATCGCTTCTGATGGGTTCCGATGTGAGTAATCCGTTGATCCCAATGAAGAC	544			
DB	418	CGGTTGCGCTCTCTGTTGACAGGTGCGATCTGAGCAAGAAGTTGATGCTGAGAAGAT	477			
QY	545	CCTAAGCTTTACCTCCACCTGGCTTTCCACAGCTACATTTATTTGCTGTGTT-TTTCAGGC	603			
DB	478	CCTAAGCTTTACCTCCACCTGGCTTTCCACAGCTACATTTATTTGCTGTGTT-TTTCAGGC	537			
QY	604	TGCTTGGGCTCTGTTTATGGTGGGTTGATCGTGGATTTTCTGTCAATCAATCAAT	663			
DB	538	TCTCTTGGAAATTTTCAGGGTTAGGGTTTCATAGTGGACTTTCTATCGCATCAACGATAGT	597			
QY	664	AGGTTTCATGGGAGGACGCCAGCGGTGTGTCTGCAGCACTAAATCGATTTCTTGG	723			
DB	598	AGGATTCATGGGAGGACGCCAGCGGTGGTGGTCTGCAACAGCTTAAAGGGTATTTTCGG	657			
QY	724	CCTTGACATTTTCAACCATCGAGCTGATATCATATCATGATCGCTCTGTTTTTCAACCA	783			
DB	658	ACTTAAACATTTTCAAGACTCTACCGATGTTAATCTGTGATGGGTTCCGTTTTTCTCCCA	717			
QY	784	AACCTATGATGGAGTGGGAAAGTCTCTGTGTAGGATGTGTCTTCAATTTTCTTCTCTCT	843			
DB	718	AACCTACGAGTGGAGATGGGAAAGTGGCGTCTCTTGGATGTGGTTTCTATTTCTTCTTCT	777			
QY	844	TAGCACAAGATACTTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTCTCAGCAATGGCGC	903			
DB	778	CTCCACAGATATTTGAGCATCAAGAAACCAAAATCTTTTGGGTGGCGGATGGCTCC	837			
QY	904	ATTGAGCTCCGTTATATTCGGAAGTCTCTTGGTTTATTTTCACTACGCCGAGAGACGG	963			
DB	838	TTTGACCTCAGTGAATCTTGGGAAGTCTCTTGGTTTACTTCTCACTCAGCGCTGAGAGATGG	897			
QY	964	TGTTGAGTATAGGAAACTGAGAGAGGTTTGAATCCACCATCACTCACAATCTGGT	1023			
DB	898	TGTTCAAGTGATAGGGGACCTGAAGAAAGGGTTGAATCCACTCTCCGGTTCTGATCTCAT	957			
QY	1024	ATTTGTGTCGCTTACATGACTACAGCTCTCAAACTGCGCATTTGCTTGGCATCATATC	1083			
DB	958	CTTTACTTCCCCTTACATGTCACAGCTGTCAAACTGCGCTCATCTGCGCATCATGTC	1017			
QY	1084	ACTTCGGAAGGAATAGCAGTAGGAAGAGCTTTTGCATGTATATAAAATTTACAAATTTGA	1143			
DB	1018	TCTCGC-GAAGGAGTAGCAGTGGGAGGAGTTTTCGATGTTTCAAGAACTACAACATAGA	1076			
QY	1144	TGSCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTCACTCTTGG	1203			
DB	1077	CGGGAACAAGAGATGATAGCGTTTGGAAATGATGAACATCATCGTTGGTTCTTCAATCTTG	1136			
QY	1204	CTACCTCAACAAGGACCAATTTTCGGGTTTCGGCTGTGAACATAACCGCTGGATGCAAGAC	1263			
DB	1137	TTACCTCAACAACCGGACCAATTTTCAAGTTCGGCAGTGAACATAACCGCGGTTGCAAGAC	1196			
QY	1264	AGCAGCTTCCAAATTAATGTCACTTCGACGTAATGTTGACATTTGATTTCTCTGACACC	1323			
DB	1197	CGCAATGTCACATAGTGTGCGGATTCGGGTTATGTTTACACTCTCTTCTCTCACACC	1256			
QY	1324	CTTGTTCATACACTCCCTGGTGTCTATCAGCTATTTATTCGATCTGCAATGCTTGG	1383			
DB	1257	GCTTTTCTACTACACCACTCGTCTGCTCTCTGCCATCATATCCGCAATGCTCGG	1316			
QY	1384	ACTCATAGATTTAAGAGCAGCCATCCATCTATTAAAGGTTGACAAATTTGACTTTGCTGT	1443			
DB	1317	ACTCATGACTATCAAGCTGCCATCCATCTCTGGAAGTTGACAAAGTTGCACTTCTCGT	1376			
QY	1444	GTGCATGAGTGCATACATTTGGCGTGGTCTTTTGGCAGTGTGAAATTTGGCTTAGTCATAGC	1503			
DB	1377	CTGCATGAGCGCTACGTTTGGGTCGTATTTCGSCAGTGTAGAGATTGGACTCGTCTGAGC	1436			
QY	1504	TATTGTAATCTGTACTTCCGGTACTTCTATTATTATTCGAAGCCCAAGGACATTCGTTTT	1563			
DB	1437	GGTGGCGATATCTATAGCGAGGTTGTTGCTGTTGTGTCGAGGCCAAAACTCGCGGTGA	1496			
QY	1564	GGSCAACATTCCAAATTTCTGTATATACCGAAATGTTGAGCACTATCAAAATGCAAAACA	1623			
DB	1497	GGAAACATACCAACACAGCATGATCTATAGGAACACTGAGCAGTACCCATCATCAAGAC	1556			
QY	1624	TGTTCTCGGAATGCTAAATTTCTAGAGATTTGATGCACCAATTTTACTTTGCCAATCCGACTA	1683			
DB	1557	CGTTCTCTGTAITCTCATCTTGAGATTGATGCTCCCATCTACTTTGCTTAATCCAGTTA	1616			
QY	1684	TTTAAGAAAGGATCACAAGGTGGATTGATGAAGAAGAAGAAGATTTAAAGCTACAGG	1743			
DB	1617	CTTGGCTGAAGAATCATTAAGGTGGATTGATGAAGAGGAGAGAGTTTAAACAATCAGG	1676			
QY	1744	GGAGACTAGTTTTCAGTATGTTATTAATGATGATGCTGTTTGGAAAAATTCATCAACAG	1803			
DB	1677	AGAGACGCTTACAAATATTAATATCTCGATATGTCAGCTGTTGGTAAATTCACACAG	1736			
QY	1804	TGGAATAAGTATGCTTTGAAGAGGTGAAGAAGATTTACAGAGAGAGAGAGCTACAGCTTGT	1863			
DB	1737	CGGTATTAGCATGATGTTGGAATTAAGAAAGTCAATTGACAGGAGAGCGTTTAAAGTTGGT	1796			



```
QY 1684 TTTAAGAAAGGATCACAGGTGGATGATGAAGAAGAAAGAAAGATTAAGCTACAGG 1743
Db 1617 CTTGCGTGAAGAAGATCATNAGGTGGATGATGAAGAAGAAAGAGAGAGTATAAACAATCAGG 1676
QY 1744 GGAGACTAGTTTTCAGGTATGTTAATTTGATATGAGTGGCTGTTGGAAACATTCGATCAAG 1803
Db 1677 AGAGAGAGCTTACAAATATATATCTCGATATGTGAGCTGTTGGTATATCGACAAG 1736
QY 1804 TGAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGAGCTACAGCTTGT 1863
Db 1737 CGGTATTAGCATGATGCTGGAATAAAGAAAGTCATTGACAGGAGAGCGTTAAAGTTGGT 1796
QY 1864 TTTGGTCAATCCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGT---TCCAAA 1920
Db 1797 ATTGTCAATTCCAAAGAGAGAGTCTGTAAGAAATTAACAGATCCAAATTCATCGGTGA 1856
QY 1921 TCATTAGGGAAGAAATGGAATCTATCTGACTGTTGAAGAGCGCGTTGGAGCATGCAACTT 1980
Db 1857 TCATTGGGCAGAGTGGATGTTCTTAACGTTAGGAGAGCAGTGGAGGCTTGTAGCTA 1916
QY 1981 CAATCTACGTGCAAGCAAAAACCAACC 2007
Db 1917 CATGCTTCACACGTTTAAACCGCAACC 1943

RESULT 7
AB004060
LOCUS AB004060.1 2037 bp mRNA linear PLN 14-FEB-2004
DEFINITION Arabidopsis thaliana mRNA for sulfate transporter, complete cds.
ACCESSION AB004060
VERSION AB004060.1 GI:2130943
SOURCE sulfate transporter.
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1
Takahashi,H., Sasakura,N., Kimura,A., Watanabe,A. and Saito,K.
Identification (20-May-1997) Hideki Takahashi, Chiba University, Faculty
of Pharmaceutical Sciences; 1-33, Yayoi-cho, Inage-ku, Chiba, Chiba
263-8522, Japan (E-mail:htaka@chiba-u.ac.jp, Tel:81-43-290-2906,
Fax:81-43-290-2905)

FEATURES
    source
        1..2037
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /db_xref="taxon:3702"
            /ecotype="Columbia"
            40..1980
                /codon_start=1
                /product="sulfate transporter"
                /protein_id="BA020282.1"
                /db_xref="GI:2130944"
            /translation="MSKRSQYHQVEIPPPQPLKSLKNTLNEILLPADPPFRIRNE
            SKTSKIEGLRHVFVILEWARGYSLEYLKSQDISGITIASLAPQGISYAQLANLPP
            ILGLYSLVPLVYATMGSSRDIAVGTAVSLLTAAMLGKEVNAVNPVKLYHLAFT
            ATFPAGLMQFCILGLRGLFVVEILSHAAIVGFMGAAATVVCLOQLKGLLGHFTST
            DIVTLRSIFSQSHMRWSGVLGCCFLIFLTKYISKKRPKLFWSAMSPLVSVIF
            GTIFLFLHDQFHGIQFIELKKGINPPSITHLVFPYPMALKVGIITGVIALRG
            IAVGRSFAMTKNTIDGNKEMIAFGMMNIIIGSSFCYLITTPFSRVAVNNAGCKTAL
            SNVMAVAVATLLFLTPFPFFYFPLVLSIIIAANMLGLVDVEAAIHLWKLDKDFPV
            CLSAYLGWVFTGTEILGLSVGISVMRLVLFVGRPKIYVYMGNIQNSEIYRNIHYPQA
            ITRSSLILHIDGPIYFANSTYLRDRGRWIDDEEDKLRISGDISQIYVLDMSAVGN
            IDPSGISMLEELNKLGRLEKLVIANPGAEVWVKLSKSTFFIESICKERIIYLTVAEV
            AACDFMLHTAKPSPVPFENNV"

CDS
```

Qy	1142	GATGGCAACAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTCACCTCT	1201
Db	1087	GATGGGAACAAGAGATGATAGCTTTTGGGATGATGAACATCTTGGTTCTTCTCTCT	1146
Qy	1202	TGCTACCTCACACAGGACCATTTTCGGTTCGGCTGTGAACTATTAACGCTGGATGAAG	1261
Db	1147	TGCTATCTCACTACCGGGCCATTTTCACGTTTCGGCGTGAACTACAAACGCGGGTGC	1206
Qy	1262	ACAGCAGCTTCCAAACATTAATAATGCTACTTGCAGTAATGTTGACATTTGTTATTCCTGACA	1321
Db	1207	ACGCGTTATCAACCGTGTGATGGGGTTCGAGTGGCGGTGACGTTACTGTTCTTAAACG	1266
Qy	1322	CCCTTGTCATTAACACTCCCTCGGTGCTATCAGCTATTATCGTATCTGCAATGCTT	1381
Db	1267	CCGTATTCTTCTACAGCCATTAGTAGTTCTGTCGTCGATCAATAATCGCTGCATGCTA	1326
Qy	1382	GGACTCATAGATTATGAGCAGCATCCATCTATTATTAAGTTGACAAAATTGACTTTGTG	1441
Db	1327	GGCTATGTGACTACGAAGCAGCCATTCATCTTTGGAACTCGACAAGTTTCGATTCTTTC	1386
Qy	1442	GTGTGCATGAGTCATACATTGCGCTGGTCTTTGGCAGTGTTCGAAATTTGCTTAGTCATA	1501
Db	1387	GTTTGTCTCTGCTACTTCTGCGGCTGTTGTTTGGTACCATCGAAATTTGTCTCATCTC	1446
Qy	1502	GCTATTGTAATATCTGTACTTCGGGTACTCTATTTATTGCAAGCGCAAGCAATTCGTT	1561
Db	1447	TCCGTGGGAATATCAGTGATGAGTTAGTTGTTCTGTCGGAAGACCAAGATTATGTA	1506
Qy	1562	TTGGGCAACATTCGAAATTCGTGTATATACCGAAATTTGGAGCACTATCAAAATGCAAAA	1621
Db	1507	ATGGGAACACATTCGAAACTCAGAAATATATAGGAACATTTGAGCATTTACCC	1566
Qy	1622	CATGTTCTCTGGAATGCTAATTTCTAGAGATTGATGCACCAATTTACTTTGCCAATGCCAG	1681
Db	1567	ACTGCTCTTCTCTTCTCATCTTCCATCTGATGGTCTTATTTACTTCGCAACTCTACC	1626
Qy	1682	TATTTAAGAGAAAGGATCACAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACA	1741
Db	1627	TACTTCGAGACAGAAATTTGNAAGTGGATCGATGAAGAGAGAAATTAATTTGAAACTAGC	1686
Qy	1742	GGGGAGACTAGTTTGCAGTATGTTTATAATTGATATGAGTCTGTTTGGAAACATTTGATACA	1801
Db	1687	GGAGATATCAGTCTACAATATATTTACTCGACATGAGTCTGTGGGGAAACATCGACAG	1746
Qy	1802	AGTGGAAATAGTATGCTTGAAGAGGTGAAGAACATTACAGAGAGAAGAGCTTACAGCTT	1861
Db	1747	AGCGGTATAGCATGCTTGGGAACCTCAACAAATCTCGGAGCTAGAGAGCTTAAGCTA	1806
Qy	1862	GTTTTGGTCAATCTGTAAAGTGAAGTGAAGAAACTGAACAAATCGAAGTTTCCAAAT	1921
Db	1807	GTGATAGCGAACCCGGGAGCAGAGGTGATGAAGAAGCTAAGCAAGTCCACCTTCAATTGAG	1866
Qy	1922	CATTTAGGAGAAATGGATCTATCTGACTGTTGAAGAGCCGTTGGAGCATGCACTTC	1981
Db	1867	AGCATTTGGTAAGAACCGAATTTATCTCACGGTGGCAGAACCGTTCGACGTTTCGATTTTC	1926
Qy	1982	AATCTACGTGCAAGCAAAACGAAACCCAAAGAAAGATGAA	2020
Db	1927	ATGCTCCACAGGCTAAACCCGACTCGCGGTACCGGAA	1965

RESULT 8  
AJ601439  
LOCUS AJ601439 2219 bp mRNA linear PLN 27-OCT-2003  
DEFINITION Brassica oleracea mRNA for sulfate transporter (l1st3.2 gene).  
ACCESSION AJ601439  
VERSION 1  
GI:37998857  
KEYWORDS l1st3.2 gene; sulfate transporter.  
SOURCE Brassica oleracea var. acephala (kale)  
ORGANISM Brassica oleracea var. acephala  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE	1	Buchner, P., Stuijver, E. E., Hawkesford, M. J. and de Kok, L. J. Analysis of the effect of sulphur nutrition and atmospheric H2S exposure on the expression of sulphate transporter genes in Curley cale	Unpublished 2 (bases 1 to 2219)
AUTHORS	Buchner, P.	Direct Submission	
TITLE	Submitted (24-OCT-2003) Buchner P., Agricultural and Environmental Division, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UNITED KINGDOM		
FEATURES	Location/Qualifiers		
source	1..2219 /organism="Brassica oleracea var. acephala" /mol_type="mRNA" /variety="acephala" /db_xref="taxon:3713" 1..2219 /gene="l1st3.2" 1..71 /gene="l1st3.2" /evidence=experimental 72..2024 /gene="l1st3.2" /function="involved in plasma membrane sulfate transport" /codon_start=1 /product="sulfate transporter" /protein_id="CAE53112.1" /db_xref="GI:37998858" /translation="MKRALQYQVETPPPPQPPKSLKNTLNEILFADDPFGKIRNES KTSKIELGRHVPFLEWARGYVILKSDVISGTTTASLAIPQISVAQLANLPPI LGLVSSFPVPMVYAIMGSSKOLAVGTAVASILLTAAMLGKESVAVENPKLYLHLAFTS TTFYSGMTCGLLRGLHVELLSHAAIIGFWAGATVVCLOQLKGLGLSHFTSTSD VVSUHSIFSQSPVWRWSGLCCFLFPITPKTISKRPKLFWSAWAPLVSVTFG SLFVFLHAQGHIOIIELEKGINPPSIHLVFTSPYVMLAKTGIITGVLAALAEGI AVGRFAMFYKNYINDGNEMAFGNMIFGFSFSCYLTTPGFSRVAENVNASCCTAVS NVMAVAVATLLFLTPLFYFPLVLSIIIAAMLGLVDYEAAMNLWRDKDFPVC LSAPGVFGTTEIGLILSVGLSVLRLLFVGRPKIYVMGNIONTPYENIROYPOAT TLSSIIILHDGPIYFANSSYLDRIGRWIDEEEEKLAKSEENSLOYIILDSAVGNI DTSGRLMEEVNKLIGRRDLKLVIANPGAELMKLSKSPFETIGKDWIHLTVAEAVS 2025..2219 /gene="l1st3.2" /evidence=experimental		
gene			
5' UTR			
CDS			
5' UTR			
ORIGIN			
Query Match	35.3%;	Score 864.8;	DB 8; Length 2219;
Best Local Similarity	66.4%;	Pred. No. 7.2e-204;	
Matches 1257;	Conservative	0; Mismatches 632;	Indels 3; Gaps 1;
Qy	125	CACCAAGTCGAGGTTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG	184
Db	96	CCCCAAGTAGAAACCCCTCCACCAACCTTCTTGAATCACTAAAGAACACTCTCAAT	155
Qy	185	GAGACTTCTTCCCTGATGACCCCTTTGAGCGAGTTTCAAGAACAA---GCCAGCTTCCAA	241
Db	156	GAGATCCCTTTGCGCGAGATCCGTTTCGAAAAATCAGGAACGAATCGAAAAACGTCGAA	215
Qy	242	AAGTTTCATGCTTGGCCTTCAGTTCCTTCTCCCATTTTTCGAATGGGCTCCCAAAATACAC	301
Db	216	AAGATCGAGTTAGGGTTGAGACACGCTGTTCCCGATCTTGGAAATGGGCTCGTGTTACAA	275
Qy	302	TTTCAGTCTTCAAGCTGACCTCATAGCTGCATCACCATCGCTAGCTTGGCCATTCCT	361
Db	276	TTGGACTATTTGAAATCAGACGTTATCTCAGGGATCAGATAGCAAGTCTTCTGCTATCTC	335
Qy	362	CAGGCGATCAGTTATGCGCAAGCTCGCCAACTCCCTCCAAATTCCTGGACTATATTCGAGC	421
Db	336	CAAGGGATTAGCTATGCTCAACTCGCTAATCTCCCTCCAATCTTGGTCTTTTACTCGAGC	395
Qy	422	TTTATACACCATTCATTTATCGGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGGACT	481

396 TTCTGCGCGCGATGGTATACCGCATTAATGGGAAGTTCAAAAGATTATAGCGGTGGGCACG 455  
482 GTGCGCGTTGGATCGCTTCTGATGGGTTTCGATGTAGTAAATGCGGTTGATGCCAATGAA 541  
456 GTAGCGGTGGCGTCTGCTGTTAAACGCGCGCAATGCTTGGAAAAGAGGTAAGCGCGGTGGAG 515  
542 GACCCAAAGCTTTACCTCCCACTGGCTTTCACAGCTACATATTATTTGCTGGTGTGTTTTTCAG 601  
516 AATCCAAACCTTTACCTTTCACCTTAGCTTTCACCTCCACTTCTTCTGCGCGGACTCATGCAA 575  
602 GCTGCTTGGCTCTGTTTAGTGTGGGCTTCATCGTGGAATTTCTGTACACATGCAACCATTA 661  
576 ACCTGCTTGGCTCTTACGGTTAGGATTTCTGTGGAGATATTGTGCGACGCGCGATC 635  
662 ATAGGGTTCAATGGAGGAGAGGAGCCACGCGGTGTGCTGAGCAACTAAAATCGATCTTT 721  
636 ATAGGGTTCAATGGCTGAGAGGAGGAGCGGTGTTGTCTCAACAACACTGAAAGGATTACTT 695  
722 GGCTTGGAGCATTTCCACCCATGGAGCTGATATCATATCAGTGATGGCTCTGTTTTCCAC 781  
696 GGCTCAGTCACCTTCCACCATTCACCTGATGTGCTGCCGTTCTTCTGTTCCATCTTTCT 755  
782 CAAACTCATAGTGGAGGAGTGGAAAGTGTGTTAGGATGTGCTTTCATTTTCTTCTCCTC 841  
756 CAGTCTCTGTGTGGAGATGGAGAGTGGGTACTAGGGTGTCTTCTTCTTCTTCTTCTT 815  
842 CTTAGCAAGAATCTTTCAGCAAAAACGACCAAGGTTTTTTTTGGGTGTGAGCAATGGCG 901  
816 CCTATACCAAGTACATTTAGCAAGAGAGACCAAGCTATTCTGGATATCTGCAATGGCT 875  
902 CCATTGACGTCGTTTATTTGGGAAGTCTCTTGGTTTTATTTCTACTCAGGCGGAGAGCAC 961  
876 CCATTGTTTCTGTCTATTTTGGGAAGCTCTTCTGTTTACTTCTCCATGCCCCAGTTCCAT 935  
962 GGTGTTCAAGTGTAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAATCTG 1021  
936 GGCATCCAAATATTAGAGAAATTAGAGAAAGGATTAATCCACGTCATCACAATCTG 995  
1022 GTATTTGTGTGCGCTTACATGACTACAGCTGTCAAACTGGCAATGTGCTTGGCATCATTA 1081  
996 GTTTTCACGTCACCTTATGTCATGCTGGCTCTTAAACCGGCATTTATCACTGGAGTCTTA 1055  
1082 TCACCTTCGGAAGATAGCAGTAGGAGAAAGCTTTGCAATGTATAAATTTACAATTT 1141  
1056 GCTCTTCTCAAGAAATTTGAGTGGGAGAAAGCTTTGCAATGTATAAAGAAATTACAACATA 1115  
1142 GATGGCAACAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTTTCACTCT 1201  
1116 GATGGTAACAAGAGATGATGATTTGGATGATGAACATTTTGGTCTCTTCTCTCT 1175  
1202 TGCTACTCAACAAGGACCATTTTTCGCTGCTGGCTGTGAACATATAACGCTGGATGCAAG 1261  
1176 TGCTATCTCACTACCGGACCGTTTTACAGTTCCGCGTGAATACAAACGCGAGTTGCAAA 1235  
1262 ACAGCAGCTTCCACATTAATATGCTACCTTGCAGTAATGTGATGTTATTTCTCTGACA 1321  
1236 ACGGCGGTATCAACCGGTGATGGCGGTTGCGAGTAGCAATGATGTTCTTAAACG 1295  
1322 CCCTTGTTCATTAACATCTCCCTGGTGGTGTCTATCATGCTATTATCGTATCTGCAATGCTT 1381  
1296 CCATATTCTTTTATACCGCGGTAGTAGTTCTGCTGCTGATCATCATCGAGCATGCTT 1355  
1382 GGACTCATAGATTATGAAGCAGCATCCATCTATTTTAAGGTTGACAAAATTTGACTTTTG 1441  
1356 GGACTCTGACTATGAAGCAGCATGAATCTTTTGGAGACTGACAAAGTTTCGATTTCTTTC 1415  
1442 GTGTGATGATGATGATACATTTGGGTGCTTTTGGCAGTGTGAAATTTGCTTAGTCATA 1501  
1416 GTATGCTCTCTGCTTCTTCTCGCGCTGTTTTTGGTACCAATTTGAGATTTGCTCATCTC 1475  
1502 GCTATTCTGATATCTGCTTCTGCTTCTGCTTCTATTTATTCGAAGGCCAAGGACATTCGTT 1561  
1476 TCTGTGGGAATATACGTCCTGAGGTACTGTTGTTCTGTTGGGAAGACCAAGAAATTTATGTA 1535

QY 1562 TTGGGCAACATTTCCAAATTTCTGTGATATACCGAAATTTTGAGCACTATCAAAATGCAAAA 1621  
Db 1536 ATGGGAAACATTTCAGAAACAGAGCCCTATAGGAAACATTTGAAACAGTACCCTCAAGCCACT 1595  
QY 1622 CATGTTCTCGAATGCTTAATTTCTAGAGATTTGATGCAACCAATTTTACTTTGCCAATGCCAGC 1681  
Db 1596 ACTCTCTCTAGTATCATCTTTCTCCATATCGATGGTCCCTATTTACTTTGCGCAACTCTTAGC 1655  
QY 1682 TATTTAAGAGAAAGGATCAACAGTGGATTTGATGAAGAAGAAAGAAATTTAAAGCTTACA 1741  
Db 1656 TACCTCGAGATAGATTTGGAAGTGGATCGATGAGGAGGAAGAAACTGAGAAAGAT 1715  
QY 1742 GGGGAGACTAGTTTGCAGTATGTTATAATTTGATATGAGTCTGTTTGGAAAACATTTGATACA 1801  
Db 1716 GAAGAAAACAGTCTACAATAACATCATCTTGAATTTGAGTCTGTGGGAAACATCGACACG 1775  
QY 1802 AGTGGAAATAGTATGCTTGAAGAGGTGAAGAGATTACAGAGAGAGAGAGACTACAGCTT 1861  
Db 1776 AGTGGTATAGAAATGCTCGAGGAAGTCAACAAATATCTGGGACGAAGAGACCTCAAGCTA 1835  
QY 1862 GTTTTGGTCAATCTCTGAAGTGAAGTGAAGAAACTGAAACAAATCGAAGTTTCCAAAT 1921  
Db 1836 GTGATAGCTAACCCAGGACAGAGTTGATGAAGAGCTAAGCAAGTCCAAAGTTCCAGAA 1895  
QY 1922 CATTTAGGGAAGAAATGGAATCTATCTGATCTGTTTGAAGAGCCCTTGGAGCATGCAACTTC 1981  
Db 1896 ACCATCGGCAAGACTGGATTCATCTCAGGTAGCAGAAAGCCCTCTCAGCCTGTGATTTTC 1955  
QY 1982 AATCTAGTGTGACGAAACAAACGACCCAAAGAA 2013  
Db 1956 GTGCTCCACAGCTAAACCGGACTCACCGGA 1987

RESULT 9  
AX53455  
LOCUS AX53455 1974 bp DNA linear PAT 22-MAR-2003  
DEFINITION Sequence 3325 from Patent WO03000898.  
ACCESSION AX53455  
VERSION AX53455.1 GI:29156269  
KEYWORDS Oryza sativa  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1  
AUTHORS Chang, H. S., Chen, W., Cooper, B., Glazebrook, J., Goff, S. A., Hou, Y. M.,  
Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.  
TITLE Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 3325 03-JAN-2003;  
Syngenta Participations AG (CH)  
FEATURES  
Location/Qualifiers  
source 1..1974  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"

ORIGIN  
Query Match 34.5%; Score 846; DB 6; Length 1974;  
Best Local Similarity 66.4%; Pred. No. 3.4e-199;  
Matches 1231; Conservative 0; Mismatches 620; Indels 3; Gaps 1;  
QY 131 GTCGAGGTTCACCGCCACAGCCGTTTTCAGTCTCTAAAGTACTCTTTGAGGAGACT 190  
Db 55 GTGCCATGCGCGCGGAGAGCCGTTCTTGGAGACGCTGGGGGGAACATGAAGAGACA 114  
QY 191 TTCTTCCCTGATGACCTTTTGAGGAGTTTCAAGAAACAAGCCAGCTTCCAAAGATTC--- 247  
Db 115 TTCTTCCGGAACGACCCGTTTCAAGGTGTCGCGGAGCGCGGTGTCGCGCGCGCG 174  
QY 248 ATGCTTGGCTTTCAGTTCTTCTTCCCATTTTTCGAATGGGCTCCCAATACACCTTTCAG 307

Db 175 GCGCGCGCGCTCCGGTACGTGTTCCGTTTCATGGAGTGGCGCGCTGTCACACCTCGGC 234  
QY 308 TTCTTGAAAGCTCACCTCTAGCTGGCATCACCATCGTAGCTTGGCCATTCTCTCAGGCG 367  
Db 235 ACCCTCAAGTCCAGCTTCATCGCGGATCACCATTGCGACGCTCGCCATCCCCAGGCG 294  
QY 368 ATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATCTTTGGACATATATTCGAGCTTTATA 427  
Db 295 ATCAGCTACGCCAAGCTCGCCAACTCCCTCCGCTCTCTCGGCTCTATTTCGAGCTTCGTG 354  
QY 428 CCACCAATGATTTATCGATGATGGTAGTCTGAGGGAATTTGCGAGTGGGAGATGTTGGCG 487  
Db 355 CCGCCGCTGCTGATCGATGATGGGAGCTCGAGGGAACCTGGCGGTGGGAGCGGTGGCG 414  
QY 488 GTTGGATCGCTCTGATGGGTTGCGATGTTGAGTAATGCCGTTGATCCCAATGAAGACCCA 547  
Db 415 GTGGCGTCTGCTGATCGGCTCGATGCTGAGGCGAGGAGTGTGCGCGCGGAGGACCCG 474  
QY 548 AAGCTTTACCTCCACCTCGCTTTTCACAGCTACATATTTGCTGGTGTGTTTTCAGGCTGCG 607  
Db 475 GCGCTGTACCTGCACTCGCGCTCAACCGCACCTTCTTCGCGGCGTGTTCAGAGCGCTG 534  
QY 608 TTGGGTCTGTTAGGTTGGGTTGATGCTGGATTTTCTGTCAATGCAACCAATATAGGG 667  
Db 535 CTGGGGGTCTTGAGGCTGGGGTTCATCTGTGACTTCTGTGCAACGCCACCATCGTCGGG 594  
QY 668 TTCATGGGAGGACGACCGGTGGTGTCTGACGCAACTAAATCGATTCTTGGCCTT 727  
Db 595 TTCTGCGGCGCGCCCGACCGGTGGTGTCTGACGAGCTCAGGGCATGTTTCGCGCTC 654  
QY 728 GAGCATTTTCAACCATGAGCTGATATCATATCAGTGTGCGTCTGTTTTTCAACCCAACT 787  
Db 655 GACCACTTCAACCAACCGCACCGACCTCGTCTCGTCTATGAGCTCGCTTCTTCCCAACG 714  
QY 788 CATGAGTGGAGTGGAAAGTCTGTGTAGATGTCTTCAATTTCTTCTCTCTTAGC 847  
Db 715 CACCTCTGCGATGGAGAGCGTCTCATGCGGCTGCGGCTTCTCTTCTCTCTCATC 774  
QY 848 ACAAGATACCTCAGCAAAAAACGACCAAGTTTTTTTTGGGTGTACAGCAATGGCGCATTG 907  
Db 775 ACCCGCTTCTTANGAAGAGGAGGCCAAGTTCTTCTGGGTATCTGCACTGCGCCATG 834  
QY 908 AGCTCGCTTATATGGGAAGTCTCTTGGTTTATTTCACTCAGCGCGAGAACGCGTGT 967  
Db 835 GCGTCTGTATCATCGGAGCGCTCTGTTGTACTTCAACATGCTGAAACCATGSCAT 894  
QY 968 GAAGTATAGGAGAACTGAGAGAGGTTTGAATCCACCATCACTCAAAATCTGTATTT 1027  
Db 895 CAAGTATTTGGTTACCTGAAGAGGCGCTGAACCCACCGTCTGCCCAAGCGCTGAAC 954  
QY 1028 GTGTGCGCTTACATGACTACAGCTGTCAAACTGGCATTTGCTTGGCATCATATCACTT 1087  
Db 955 TGTGCGCGTACATGATGCTGGCCCTGAGACCGGGATCATCACTGGCGTCAATGCGCCCT 1014  
QY 1088 GCGGAAGGAATAGCAGTAGGAAGCTTTGCAATGTATAAAATTTACAATATTTGATGCG 1147  
Db 1015 GCTGAGGGGATGCGCTAGGAGGAGCTTTGCAATGTTCAAGAAATTTACACATCGACGA 1074  
QY 1148 AACAAAGAGATGATAGCTATTGGGACCATGAACTGATTTGGTTCTTTACCTCTTCTAC 1207  
Db 1075 AACAAAGGATGATGCGGTTTGGAAAGATGAACATTTGGGGATCACTCACTCTTCTGTAC 1134  
QY 1208 CTCACAAAGGACCAATTTTTCGCGTTCGGGTGTGAATATAAAGCTGGATGCAAGACGCA 1267  
Db 1135 CTCACACCGGCTCGTCTGAGGTGCGCGCTCACTCAACGCGGCTGCAAGCGGG 1194  
QY 1268 GCTTCCAAATTAATGTCATTTGAGTAATGTTGACATTTGTTATTCCTGACACCTTTG 1327  
Db 1195 ATGTCCAAAGTATCATGTCTGGTGGCGGTGATGATCAGCTGCTGTTCTTTCGCGCGCTG 1254  
QY 1328 TTCCATTACATCCCTCTGGTGGTATCAGCTATTTATGCTATCTGCAATGCTTGGACTC 1387  
Db 1255 TTCCACTACACGCGCTGGTGGTGTCTGCGGATCATCATGTGCGCGGATGCTGGGGCTG 1314

QY 1388 ATAGATTATGAAGCAGCCATCCATCTATTATTAAGTTTGACAAATTTGACTTTTGTGGTGTG 1447  
Db 1315 ATCGACTACCCGCGCGCTCCACCTGTGGAGGTGGACAGAGTGGACTTTCTGCGTCTGC 1374  
QY 1448 ATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTACTATAGTATT 1507  
Db 1375 CTCGGCGCTACCTCGGCGTCTCTTCGGCAGCGTCGAGATCGGCTCGTCTGCGCGTC 1434  
QY 1508 GTAATATCTGACTTTCGGGTACTTCTATTTATTGCAAGGCCAAGGACATTCGTTTGGGC 1567  
Db 1435 GGGATCTCCATCTCCGCGTCTGCTGTTTCGGCGCGCGCGAGGACGACGCTGCTCGGG 1494  
QY 1568 AACATTTCCAAATTTCTGTGATATACCGAATGTTGAGCACTATCAAAATCCAAAACATGTT 1627  
Db 1495 AACATCCCAACACATGATCTACCGCGGATGGACAGTACACCGCGCGCAGAGGTC 1554  
QY 1628 CTTGGAATGCTAAATCTAGAGATTGATGCACCAATTTTACTTTGCCAATGCCAGCTATTTA 1687  
Db 1555 CCGCGCTGCTGCTCGCGCTCGACTCGGCCATCTACTTCAACAAACCCAGCTACCTG 1614  
QY 1688 AGAGAAAGGATCACAGGTGGATTGATGAAAGAAAGAAATTTAAAGCTACAGGGAG 1747  
Db 1615 CGTGAGAGATCGCGCGTGGATCGACGACGAGGAGGACCAAGTGCAGAGGAGAGGCGAG 1674  
QY 1748 ACTAGTTTGCAGTATGTTTATTAATTTGATATGATGAGTGTGTTGGAAACATTCATCAAGTGA 1807  
Db 1675 ATGGGCATTTCACTGCTGCTCTGACATGGGTGCGTTCGGAAGCATCGACCCAGCGGG 1734  
QY 1808 ATAAGTATGCTTGAAGAGTGAAGAGATTACAGAGAGAGAGACTACAGCTTGTGTTT 1867  
Db 1735 ACGAGCATGCTGATGAACTCAGGAAGACCTCGACAGGAGGCGCTTCAGATTGTGTTG 1794  
QY 1868 GTCATCTCTGAAGTGAAGTGAAGAACTGAAACAAATCGAACATTCGAAATCATTTA 1927  
Db 1795 GCGAACCCTGGGAGCGAGATCATGAAGAGATTGGACAGCTCCAGGTGCTTGAGGCGATC 1854  
QY 1928 GGGAGAAATGATCTATCTGACTGTGAAAGAGCGCTTGGAGCATGCAACTTC 1981  
Db 1855 GGCATGAGTGGATCTTCCCAAGTGGCGAGGCGGTGGCGGATGCGCATTC 1908

RESULT 10  
AK066932  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

AK066932 2327 bp mRNA linear PLN 24-JUL-2003  
Oryza sativa (japonica cultivar-group) cDNA clone:J013093F02, full insert sequence.  
AK066932  
FLI CDNA; CAP trapper.  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

1  
The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team.; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Naniwa,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group.; Oono,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurokawa,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niihara,J., Ikeda,T., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN., Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Yoshino,M. and Hayaishizaki,Y.



TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice	
JOURNAL	Science 301 (5631), 376-379 (2003)	
MEDLINE	22752273	
PUBMED	12869764	
REFERENCE	2 (bases 1 to 2327)	
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Ikeda, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanaoka, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Ootani, N., Oka, Y., Ootani, Y., Ootani, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.	
TITLE	Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp)	
JOURNAL	Tel: 81-29-838-7007, Fax: 81-29-838-7007	
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice. URL : <a href="http://cdna01.dna.affrc.go.jp/cDNA/">http://cdna01.dna.affrc.go.jp/cDNA/</a> NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M. FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryo, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanaoka, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ootani, N., Oca, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashizaki, Y. Location/Qualifiers 1. 2327 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /culturvar="Nipponbare" /db_xref="taxon:39947" /clone="J013093F02"	
FEATURES	Source	
ORIGIN	Query Match 34.5%; Score 846; DB 8; Length 2327; Best Local Similarity 66.4%; Pred. No. 3.4e-199; Matches 1231; Conservative 0; Mismatches 620; Indels 3; Gaps 1;	

QY	131	GTGAGGTTCCACCGCCACAGCGTCTTCAAGTCTCTAAGTACTCTTTGAAGGAGACT	190
DB	166	GTGCCGATGCGCGCGGGAAGCGTCTCTGAGACGCTGGGGGGAACATGAAGGAGACA	225
QY	191	TTCTTCCCTGATGACACCTTTGAGCGAGTTCAAGAAACAGCCAGCTTCCCAAGATTTC	247
DB	226	TTCTTCCCGGACACCGTTCAGGGTGGTGGCGGAGCGCGGTGCGGGCGCGCGCG	285
QY	248	ATGCTTTGGGCTTCAGTCTCTTCTTCCCAATTTTCCGATGGGCTCCCAATATACACTTT	307
DB	286	CGCGCGCGCTCGGTAAGTGTTCCTGTCATGAGTGGGCGCGTGTACACCTCGGC	345
QY	308	TTCTTCAAAAGCTGACCTCATATAGCTGCATCAACATCGTAGCTTGGCCATTTCTC	367
DB	346	ACCTCAAGTCCGACCTCATCGCGGCATCACCATTTGCCAGCTCGGCATCCCCAGGC	405
QY	368	ATCAGTTATGCCAAGCTCGCCAACTCCCTCCATTTCTTGGACTATATTCGAGCTTT	427
DB	406	ATCAGCTACGCCAAGCTCGCCAACTCCCTCCGCTCTCGGCTCTATTTCGAGCTTC	465
QY	428	CCACCATTTGATTTATCGCATGATGGGTAGCTCGAGGATTTGGCAGTGGGACTGTGG	487
DB	466	CGCGCGCTGTGTACGCGATGATGGGAGCTCGAGGAGCTGCGCGTGGGACGGTGG	525
QY	488	GTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATCCCGTTGATCCCAATGAAGCC	547
DB	526	GTGGCGTCTGCTGATCGGTCGATGCTGAGCGAGGAGTGTGGCGCGGAGGACCG	585
QY	548	AGCTTTTACCTCCACCTGCTTTTACAGCTACATATTATTTGCTGGTCTTTTTCAGG	607
DB	586	GCGCTGATCGACCTCGCGCTCACCGCCACCTTCTTCCGCGCGTGTTCAGAGGCTG	645
QY	608	TTGGGCTGTGTAGGTTGGGTTGATCGTGGATTTTCTGTACATGCAACATAATAGG	667
DB	646	CTGGGGCTTTGAGGCTGGGTTTCTCTGGAATCTCTGTGCAACCGCACCATCTCG	705
QY	668	TTCATGGGAGGAGCAGCCAGGTGTGTGTGACAGCAATAAAATCGATTTCTTGGCT	727
DB	706	TTCATGGGCGGCGCCCAACCGTGTGTGCTGTCAGCAGCTCAAGGCGCATGTTTGG	765
QY	728	GAGCATTTACCCATGAGCTGATCATCATCATGATGCTGCTCTTTTTCACCCAACT	787
DB	766	GACCACTTACCAACCGCCACCGACCTCGTCTCGTATGAGCTCCGCTTCTCCCAAC	825
QY	788	CATGAGTGGAGTGGGAAAGTGTGTGTAGGATGTGTCTCATTTTCTTCTCTCTC	847
DB	826	CACCTCTGGGATGGGAGAGCTGCTCATGGGCTGGGCTTCTCTTCTCTCTCTC	885
QY	848	ACAAGATCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGCAT	907
DB	886	ACCGCTTCTTCAGCAAGAGGAGGCGCAAGGTTCTTCTGGGTATCTGACGCTGGCC	945
QY	908	AGTCTGTTATTTGGGAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGAGCGGTGT	967
DB	946	GCGTGTGTCATCATCGGAGCGCTCTCGTGTGTACCTCACATGCTGAAACCATG	1005
QY	968	GAAATGATAGGAGAACTGAAGAAGGTTTGAATCCACCATCATCTCACAATCTGGT	1027
DB	1006	CAAGTGTGTTTACTCTGAAAGAGGCGCTGAAACCCACCGTCTGCCAACGCTG	1065
QY	1028	GTGTCGCTTACATGATGATGAGCTGTCAAACTGGCATTTGTGTTGGCATCATAT	1087
DB	1066	TGTCGCGCTGATGATGCTGGCCCTGAAGACCGGATCATCACTGGCGTCAATGG	1125
QY	1088	CGGAGGATAGCAGTAGAGAGAGCTTTGCATGATATAAATAATATATTTGATGG	1147
DB	1126	GCTGAGGAGATTGCCGTAGGAGGAGCTTTGCAATGTTTCAAGAAATATACCA	1185
QY	1148	AACAAAGAGATGATAGCTATTGGGAGCATGAAGTGTGTTTCTTTTCACTCTTGT	1207
DB	1186	AACAGGAGATGATCGGTTTGGACGATGAACATTTGGGATCACTCACTCTTGT	1245

```
QY 1208 CTCAACACAGGACCAATTTTCGGTTCGGCTGTGAACATATAACGCTGGATGCAAGACAGCA 1267
Db 1246 CTTACACACCGTCCGTTCTCGAGGTTCGGCGTCAACTACAAAGCCGGCTGCAAGACGGCG 1305
QY 1268 GCTTCCAAATTAATATGTCATCTTGCAAGTAATGTTGACATTTGATTCCTGACACCCCTTG 1327
Db 1306 ATGTCGAACGTGATCATGTGCGGTGGCGGTGATGATCAGCGTGTGTTTTCGACCGCGCTG 1365
QY 1328 TTCCATTTACATCCCTCGGTGGTGTATCAGCTATTATTCATCTGCAATGCTTGGACTC 1387
Db 1366 TTCCACTACACGCGTGTGTTGCTGTGCGCGATCATCATGTCGGGATGCTGGGGCTG 1425
QY 1388 ATAGATTAAGAACAGACCAATCCATCTATTTAAGGTTGACAAATTTGACATTTGTTGTGC 1447
Db 1426 ATCGACTACCCGCGCGCTCCACCTGTGGCAGGTGACAAAGGTGACTTCTCGCTCTGC 1485
QY 1448 ATCAGTGCATACATTTGGCGTGTCTTGGCAGTGTGAAATTTGGCTTAGTCATAGCTATT 1507
Db 1486 CTCGGCGCCTACCTCGCGGTGCTCTTGGCAGCGTCCAGATCGGATCGGCTCGTCCGCTC 1545
QY 1508 GTAATATCTGACTTCGGGTACTTCTATTATTATTGCAAGGCCAAGGACATTTGTTTGGGC 1567
Db 1546 GGGATCTCCATCTTCGCGTGTCTGTGCTGCGCGCCGAGGAGCAGCGTGTCTGGG 1605
QY 1568 AACATTCCAAATTTCTGTGATATACGAAATGTTGAGCACTATCAAAATGCAAAACATGTT 1627
Db 1606 AACATCCCAAACACGATGATCTACCGCGGATGGACAGTACACCGCGCGCAGAGGGTC 1665
QY 1628 CTTGGAAATGCTAATTTCTAGAGATTGATGCACCAATTTACTTTGCCAATGCCAGCTATTTA 1687
Db 1666 CCGGCGGTGCTGCTGTCTCGCGTCTGCTGCTGCGCGCCGAGGAGCAGCGTCTACCTG 1725
QY 1688 AGAAGAGGATCAAAAGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACAGGGGAG 1747
Db 1726 CGTGAGAGAGTCCGCGGTGGATCGACGACGAGAGGACCAAGTGCAAGAGAGGGCGAG 1785
QY 1748 ACTAGTTTGCAGPATGTTATTAATTTGATATGAGTGTCTTTGGAAACATTTGATACAACTGA 1807
Db 1786 ATGGGCAATTCAGTACGTGCTCGACATGGGTGCGTGGAGAGCATCGACACCGCGG 1845
QY 1808 ATAGATGCTTGAAGAGGTGAAGAGATTCAGAGAGAGAGAGCTACAGCTTGTGTTTG 1867
Db 1846 ACAGCATGCTGGATGAATTCAGGAAGACCCCTGGACAGGAGGGGCTTCAGATTTGTGTTG 1905
QY 1868 GTCAATCTCTTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCAAATCATTTA 1927
Db 1906 GCGAACCCGGGAGCGAGATCATGAAGAGTTGGACAGCTCCAGGTGCTTGAGGCGATC 1965
QY 1928 GGAAGAAATGATCTATCTGACTGTGTAAGAGCGGCTTGGAGCATGCAACTTC 1981
Db 1966 GGCCATGAGTGGATCTTCCCACGGTGGCGAGCGGTGGCGAGTGGCAGCTTC 2019
```

```
RESULT 11
LOCUS AR477457
DEFINITION Sequence 3 from patent US 6696292.
ACCESSION AR477457
VERSION AR477457.1 GI:47234915
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1981)
AUTHORS Allen, S.M., Falco, S.C. and Thorpe, C.J.
TITLE Genes encoding sulfate assimilation proteins
JOURNAL Patent: US 6696292-A 3 24-FEB-2004;
FEATURES
    Location/Qualifiers
        1..1981
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
```

```
Query Match 31.2%; Score 765.2; DB 6; Length 1981;
Best Local Similarity 65.8%; Pred. No. 4.3e-179;
Matches 1112; Conservative 0; Mismatches 578; Indels 0; Gaps 0;

QY 313 GAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTTGGCCATTCCTCAGGGGATCAG 372
Db 7 GGAGTCGACCTGATCGCGGCATCACCATCGCCAGCTTCGCCATCCCGCAGGCAATCAG 66
QY 373 TTATGCCAAGCTCGCCAACTCCCTCCCAATTTCTTGAGCTATATTGAGCTTTATACCAACC 432
Db 67 CTACGCCAAGCTCGCCAACTCGCGCCGCTGCTCGACTCTACTCGAGCTTCTGTCGGCTGC 186
QY 433 ATTGATTTATGCGATCATGGGTAGCTCGAGGGAATTTGCGAGTGGGACTGTGCGCGTGG 492
Db 127 GCTGGTGTACGCGCTGATGGGAGCTTCAAGACCTGGCGGTGGGACGGTGGCGGTGCG 186
QY 493 ATCGCTTCTGATGGGTTCGATGTTGAGTAAATGCGGTTGATCCCAATGAAGACCCAAAGCT 552
Db 187 GTCGCTGCTCATCAGCTCCATGCTCGGCAGCGAGGTGTCGCCGACGGAGAACCCCGTGT 246
QY 553 TTACCTCCACCTGGGCTTTTCCAGCTTACATTTATTTGCTGGTGTGTTTTCAGGGCTGCTTGG 612
Db 247 CTAACCTGACCTCGCTTCCAGCCACCTTCTTGCGCGGCGTCTTCCAGGCTTGGTCTGG 306
QY 613 TCTGTTTAGTGGGTGATCGTGATTTTCTGTCAATGCAACCATTAATAGGGTTCTAT 672
Db 307 CTTCTCAGTGGGCTTTCATCGTGGACCTGCTGCGACGCGACGATCGTGGGGTTCAT 366
QY 673 GGGAGGACGACCGGTGGTGTCTGCGACCACTAAATCGATTCCTTGGGCTTGGAGCA 732
Db 367 GCGCGCGCGGCGACGCGTGTGCTGCGACGAGCTGAAGGGCATGCTGGGCTTCGTCCA 426
QY 733 TTTTACCCATGAGCTGATATCATATCAGTGATCGCTCTGTGTTTTCACCCAACTCATCA 792
Db 427 CTTTACACCTCCACCGAGCTGCTCTCGTCAATGAATCCGTTTCAGCAGACACACCA 486
QY 793 GTGGAGTGGGAAAGTGTGTTAGGATGTGTTCTCAATTTTCTTCTCTCTTAGCAAG 852
Db 487 GTGGCGTGGGAGAGCGTCTCTCGGCTGGGCTTCTCTTCTTCTCTCTCTCTCAACCG 546
QY 853 ATACTTCAGCAAAAACGACCAAGTTTTCCTGAGTGTGAGCAATGGGCGCATTTGACGTC 912
Db 547 CTTTATCAGCAAGAGCGCTCCCAAGCTGTTCTGAGATCTCCGCGCGCGCGCTTGAAGTC 606
QY 913 CGTTATATTGGGAAGTCTCTCGTGTATTCTTCACTCAGCGCGAGAGACGCGTGTGAAAGT 972
Db 607 CGTGTGCTCGGAGGGTCTGCTGTA CTTCAAGCAGCGTGAAACCAAGCATCGAAGT 666
QY 973 GATAGAGAACTGAAGAAGGGTTTGAATCCACCATCATCAAAATCTGGTATTTGTGTGTC 1032
Db 667 GATCGGTTACCTGAAGAAGGCTGAATCCACCGTGGTGAACAAGCTTCAATTTCTCAC 726
QY 1033 GCCTTACATGACTACAGCTGTCAAAACTGGCATTTGCTGGCATCATCATCTTCCGA 1092
Db 727 GCGCTACATGATGCTCGGCGCTCAAGACTGGGATCATCACCGGCGCTATTGGCCTTCGCG 786
QY 1093 AGGAATAGCAGTAGGAAGAAGCTTTTGCATGTATATAAAATTTACAATATTGATGGCAACA 1152
Db 787 AGGAATCGCGCTGGGAGGAGCTTCGCCATGTTCAAGAACTACCAATGACGAGCAACA 846
QY 1153 AGAGATGATAGCTATGGGACCATGAAGTGTGCTTCTTTTCACTCTTGTGCTACTCTAC 1212
Db 847 GGAGATGATCGGATCGGACCGATGAAGTCTCTGGGCTCGCTCAGGTGTGCTACTCTGAC 906
QY 1213 AACAGGACCAATTTTCGGCTTGGCTGTAACATATAACGCTGGATGACAGACGAGCTTC 1272
Db 907 CACGGGGCCCTTCTCGCGCTCGCGCTGAACATACAAACGCGGCTGAGACGCGCATGTC 966
QY 1273 CAACATTAATGTCACCTTGCAGTAATGTTGACATTTGTTATTTCTTGACACCCCTTTTCCA 1332
Db 967 GAACGTGGTCACTGCTGGCGGTGATGGTCAAGCTGCTGTTTCTTGACCGCGCTGTCCA 1026
QY 1333 TTACACTCCCTGGTGGTGTATCATGCTATTATCGTATCTGCAATGCTTGGATGCTATAGA 1392
```



Db 1027 CTACACGCCGTGGTGGTCTGTCGGCATCATCGTCTCCGCGATGCTGGGCGCTGGTCGA 1086  
Qy 1393 TTATGAAGCAGCCATCCATCTATTAAAGGTTGACAAATTTGACTTTTGGTGTGTCATGAG 1452  
Db 1087 CTTGGGCGCGCGCTGCACCTGTGGCGGTTCGACAAAGGTTCGACTTCTGCGTCTGGCGCG 1146  
Qy 1453 TGCATACATTTGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCTATTGTAAT 1512  
Db 1147 CGCGTACCTGGCGTGGTCTTTGGCGAGCTTCGAGTTCGGCGTGGTGGTGGTGGTGGTGG 1206  
Qy 1513 ATCTGTACTTCGGGTACTTCTATTATTATGCAAGGCCAAGGACATTCGTTTTGGGCAACAT 1572  
Db 1207 CTCCTGTCTGGCGTCTGCTGTTGTCGCCGCGCCAGGACCGAGTGTCTGGCAACAT 1266  
Qy 1573 TCCAAATTCGTGATATACCGAAATGTTGACACTATCAAAATGCAAAACATGTTCTCTGG 1632  
Db 1267 CCCGGCACCATGTGTACCGGAGATGGACCACTAGCCCGCGCGAGCGGTGGCGCG 1326  
Qy 1633 AATGCTAATTCATAGATTTGATGACCAATTTACTTTGCCAATGCCAGCTATTAAAGAGA 1692  
Db 1327 CGTCTGTGTGGCGTTCGACGCGCGCGCTACTTTCGCCAACGCCAGCTACCTGCGAGA 1386  
Qy 1693 AAGATCAACAGGTGGATTGATGAAGAGAGAGAGAAATTAAGCTACAGGGAGACTAG 1752  
Db 1387 GAGATCTCGCGGTGGATTCGACGAGAGAGAGAGCGCACCAAGAGCCAGGGCGAGATGG 1446  
Qy 1753 TTTGAGTATGTTAATTTGATATGATGCTGTGTTGGAACATTTGATCAAGTGGAATAAG 1812  
Db 1447 CGTGGGTACGTTCTCTCGACATGGGTGCCATCGGTAGCATCGACAGCGGGAGCGAG 1506  
Qy 1813 TATGCTTGAAGAGTGAAGAGATATACAGAGAGAGAGCTACAGCTTGTGTTTGGTCAA 1872  
Db 1507 CATCTGACAGCTCAACAGTCTCTTGGACAGAGGGAATGACATCGTGTGGCGAA 1566  
Qy 1873 TCCTGTAAGTGAATGATGAGAACTCAACAAATCGAAGTTCCAAATCAATTTAGGGAA 1932  
Db 1567 CCCGGCAGCAGATCATGAAGAAGCTGGACAGCTCCAAGGTGCTGGAGCAGATCGGCCA 1626  
Qy 1933 GAAATGGATCTATCTGACTGTTGAAGAGCGGTGGAGCATGCAACTTTCAATCTACGTGC 1992  
Db 1627 CGAGTGGGTGTTCCGAGCGTGGCGAGCGGTGGCGCTGTCGCACTACGTGTCGACTC 1686  
Qy 1993 AAGCAAAACG 2002  
Db 1687 GCACAAGCCG 1696

RESULT 12  
AR477463  
LOCUS AR477463 2067 bp DNA linear PAT 14-MAY-2004  
DEFINITION Sequence 15 from patent US 6696292.  
ACCESSION AR477463  
VERSION AR477463.1 GI:47234921  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2067)  
AUTHORS Allen,S.M., Falco,S.C. and Thorpe,C.J.  
TITLE Genes encoding sulfate assimilation proteins  
JOURNAL Patent: US 6696292-A 15 24-FEB-2004;  
FEATURES  
source 1. .2067  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 29.4%; Score 720.6; DB 6; Length 2067;  
Best Local Similarity 62.2%; Pred. No. 5.3e-168;  
Matches 1134; Conservative 0; Mismatches 689; Indels 0; Gaps 0;  
Qy 143 CCGCCACAGCGGTTTTCAAGTCTCTAAAGTACTCTTTTGAAGGAGACTTTCTTCCCTGAT 202

Db 8 CCACACAGACCACTCCCAAACTTAGGCACAGAGTCTCCGAAATCTTCTTCCAGAT 67  
Qy 203 GACCCCTTTGAGCGAGTTCAAGAACACAGCCAGCTTCCAAAGAGTTCATGCTTGGGCTTCAG 262  
Db 68 GACCCCTCCACCGTTTCAAGAACCAAACTCGCTTTAAAGAGTTCCTCTCGACTTCAG 127  
Qy 263 TCTCTTCTCCCAATTTTCGAATGGGCTCCCAATACACCTTTTCAGTGTCTTGAAGAGCTGAC 322  
Db 128 TATCTCTTCCCAATTTTCGACTGGGCGCCAAATCTACAATCTTACCCCTCTCCGCTCTGAC 187  
Qy 323 CTATAGCTGCGCATCACATCGCTAGCTTGGCCATCTCTCAGGCGCATCAGTTATGCCAAG 382  
Db 188 CTATCTCTGGCCTCACATTCGACGCTCGCCATCTCTCAGGGAATCAGTTATGCCAAG 247  
Qy 383 CTGCGCAACCTCCCTCCAAATCTTGGACTATATTCGAGCTTTATACCAACCATTCATTTAT 442  
Db 248 CTTGCCAACTTGCACCTATTTCTTGGATATATTCGAGTTTGTTCCTCCCATTCATATAC 307  
Qy 443 GCGATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGACTGTGGCGGTGGATCGCTTCTG 502  
Db 308 TCGCTGTCTGAAAGTTCTAGACATCTTGGTGTGGACCTGTTCCATTGGCTCTTTGGTC 367  
Qy 503 ATGGGTTGATGTTGAGTATGCGGTGATCCCAATGAAGACCCAAAGCTTTACCTCCAC 562  
Db 368 ATGGGATCAATGTTAAGTGATAAATTTCTTACACTCAAGAACCTATTTCTCTATCTGGGA 427  
Qy 563 CTGGCTTTACAGCTACATTTATTTGCTGGTGTGTTTTTTCAGGCTGCTTGGGTCTGTTTGG 622  
Db 428 TTGGCTTTACCGCCACTTCTTTGCTGGTGTATTCGAGTTTCTCTGGGTATATTAAAG 487  
Qy 623 TTGGGTTGATCGTGGATTTTCTGCATGCAATGCAACCATATATAGGTTTCATGGGAGGACA 682  
Db 488 CTAGGCTTCGTAATTTGATTTTCTGTCGAGGCAACGCTGGTTGGATTCACAGGCGGTGCT 547  
Qy 683 GCGACGGTGGTGTCTGCGAGCAACTAAATTCGATTTCTGGCCCTTGAGCAATTCACCCAT 742  
Db 548 GCCATATTGTTGTCACTGCGACGAGCTGAAAGGTTTACTTTGGAATAGTGCACTTTTACCAGC 607  
Qy 743 GGAGCTGATATCATATCATGATGATGCTGTTTTCACCCAACTCATATGATGAGGAGTG 802  
Db 608 AAGATGCAATAATTCAGTAAACGATCTCTGTTTTCAAGCAAGACACGAGTGTCTATGG 667  
Qy 803 GAAAGTGTGTGTAGGATGTCTTCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAGC 862  
Db 668 CAAACCATTTCTTTTGGGATTCGGCTTCTCTGGTCTTCTTGGCTGACAAAGGCACTTAGT 727  
Qy 863 AAAAAAGCAGGATTTTGGGTGTGAGCAATGGCGCATTCAGCTCCGTATATTG 922  
Db 728 TTGAGGAAACCAAAACTATTTCTGGGTTTCAGCAGCTGCCCCCATTCACATCAGTTATTCTG 787  
Qy 923 GGAAGTCTCTTGGTATTATTTCACCTCAGCGGAGACGCGTGTGGAAGTGATAGGAGAA 982  
Db 788 TCAACCATTTAGTCTTTCTCTGGAATAAGACTCATCAAAATTCAGTTATTTGGGCAC 847  
Qy 983 CTGAAGAAGGTTTGAATCCACCATCACTCACAATCTGGTATTTGTGTGCGCTTACATG 1042  
Db 848 TTACCAAAAGGAGTTAATCCACCATCAGCAAAACATGTTATCTTCAATGGTCTTACTTG 907  
Qy 1043 ACTAGCTGTGCAAACTGCGCATTTGCTTGGCATCATATCATCTTGGGAGGAATAGCA 1102  
Db 908 GGTCTTGTCTATCAAACTGCGCATCATCAGGGGATCTTATCTCTACTGGAAGGAATGCA 967  
Qy 1103 GTAGGAAGAAGCTTTGCAATGTATAAAATTAACATATTGATGGCAACAAAGAGATGATA 1162  
Db 968 GTAGGAGAACATTTGCTCTCACTTAAGAACTACAGGTGGATGGAACAAAGAAATGATG 1027  
Qy 1163 GCTATTGGGACCATGAACGCTAGTTGTTGTTCTTTTCACTCTTGTCTACCTCACAACAGGACCA 1222  
Db 1028 GCCATTTGGTCTAATGAACATAGCTGGCTGCTGTTCTTCTATGTTATGTTACACGGGATCC 1087  
Qy 1223 TTTTCGCGTTCGGCTGTGAACCTATAACGCTGGATGCAAGACAGAGCTTCCCAACATTATA 1282

Db	1088	TTTTCTCGATCGGCTGTAACTATAATGCTGGAGCACAGACAACAGTTCCTCAATATAATC	1147
Qy	1283	ATGTCACCTTGACGTAATGTTGACATTTGTAATTCCTGACACACCTGTTTCCATTACACATCCC	1342
Db	1148	ATGCTCAGCTGTTAGTGACACTTCTGTTTCTCATGCTCTTTTCTACTATACACCA	1207
Qy	1343	CTGGTGTGTATCAGCTATTATCGTATCTGCAATGCTTGGACTCATAGATTATGAAGCA	1402
Db	1208	AATGTTGCTTAGCGGCCATTATCATCTGCTGTGATTGGTCTAATAGATTATCAATCT	1267
Qy	1403	GCCATCCATCTATTAGGTTGACAAATTTGACATTTGCTGTGTGTCATGAGTCATACATT	1462
Db	1268	GCATATAAATTTGGGAAGTTGACAAACCTTGAATTTCTTGGCCCTGTTTGTCTCCTTTTTT	1327
Qy	1463	GGCGTGTCTTTGGCAGCTGTGAAATTTGGCTTAGTCTAGTCTATTGTAATCTGTACTT	1522
Db	1328	GGGTTCTGTTTCACTTTCAGTGCCTTTAGTCTTGGTATAGCGGTTATCAATCAGTCCTC	1387
Qy	1523	CGGTAATCTTCTATTATTGCAAGGCCAAGGACATTCGTTTGGGCAACATTTCCAAATTC	1582
Db	1388	AAGATCCTGCTTCTATGTCACCTCGACCAACACACTTTGGTTTGGGAATATACACGAACA	1447
Qy	1583	GTGATATACCAATGTTGAGCACTATCAAAATGCAAAACATGTTCTCGAATGCTAAAT	1642
Db	1448	CAAAATATCCACAACTAAACCAATACAAAAAGCTTTTAAGAGTTCTCTTCAATTTCTCAT	1507
Qy	1643	CTAGAGATTGATGACCAATTTTACTTTGCCAATGCCAGCTATTATTAAGAGAAAGGATCACA	1702
Db	1508	TTGGCTGTGAGTCTCAATCTATTTTGCTTAACCTACTTATCTTCAAGAAAGGATACTG	1567
Qy	1703	AGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACAGGGGAGACTAGTTTTCAGTAT	1762
Db	1568	AGATGGGTTGCGAAGAGGAAGAGCATATAAAGCTTAATAATGGAGCTCCCATTTGAAGTC	1627
Qy	1763	GTATATAATTGATGCTGCTGTGTGGAACATTTGATACAGTGGATAGTATGCTTGAA	1822
Db	1628	ATAATTTTAGACATGACAGCTGTACAGCCACAGACAACAGTGGGCTTGACACTTTATGT	1687
Qy	1823	GAGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTTGTGTTGGTCAATCTGTAAGT	1882
Db	1688	GAACCTAGAAGATGCTGGAGAGAGATCACTTGAAGTTTGTGCTGGCAATCTCTTTGGA	1747
Qy	1883	GAAGTATGAAGAACTGAACAAATCGAAGTTCCAAATCATTTAGGGAAGAAATGGATC	1942
Db	1748	AATGTGATGAAAAAATGTCATAAGTCAACATTTTGGATTCTTTTGGATTAAAAAGGAGTC	1807
Qy	1943	TATCTGACTGTTGAAGAGGCCGT	1965
Db	1808	TATCTCAGTGGGAGAAGCTGT	1830

RESULT 13	
AJ704374	
LOCUS	AJ704374
DEFINITION	Brassica oleracea var. acephala mRNA for plasma membrane sulphate transporter (ST3.4 gene).
ACCESSION	AJ704374
VERSION	1
KEYWORDS	plasma membrane sulphate transporter; ST3.4 gene.
SOURCE	Brassica oleracea var. acephala (kale)
ORGANISM	Brassica oleracea var. acephala Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE	1
AUTHORS	Buchner,P., Stuiver,E.E., Hawkesford,M.J. and de Kok,L.J.
TITLE	Analysis of the effect of sulphur nutrition and atmospheric H2S exposure on the expression of sulphate transporter genes in Curley cale
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2299)
AUTHORS	Buchner,P.
TITLE	Direct Submission

JOURNAL	Submitted (06-MAY-2004) Buchner P., Crop Performance Improvement, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UNITED KINGDOM
FEATURES	
source	Location/Qualifiers 1..2299 /organism="Brassica oleracea var. acephala" /mol_type="mRNA" /variety="acephala" /db_xref="taxon:3713"
gene	1..2299 /gene="ST3.4"
5' UTR	1..138 /gene="ST3.4" /evidence=experimental
CDS	139..2109 /gene="ST3.4" /codon_start=1 /product="plasma membrane sulphate transporter" /protein_id="CAG28416.1" /db_xref="GI:47109356"
3' UTR	2110..2299 /gene="ST3.4" /evidence=experimental
ORIGIN	
Query Match	27.8%; Score 681; DB 8; Length 2299;
Best Local Similarity	61.0%; Pred. No. 3.8e-158;
Matches 1104; Conservative	0; Mismatches 705; Indels 0; Gaps 0;
Qy	158 TTCAAGTCTCTAAAGTACTCTTTTGAAGGAGAGACTTTCTCCCTGATGACCTTTGAGGCAG 217
Db	259 TTTTCAGAACTCAAAAAGCGTTTGGTGATGTTTCTCCCTGATGATCCTTTAGAGAGG 318
Qy	218 TTCAAGAACCAAGCAGCTTCCAAGAGTTCATGCTTGGCCTTCAGTTCCTTCTCCCAT 277
Db	319 TTAGAAACACAGATGGAGAAACAAAGTCAITCTTGGTCTACAAAGCTTGTTCCTATA 378
Qy	278 TTCTGAATGGCTCCCAAAATACACACTTTCAGTTCCTTGAAAGCTGACCTCATAGCTGGCATC 337
Db	379 TTCCCATGGGTTCTCAGTACGATCTCAAGCTTTTTAGTCTGATGTTATCTTGGACTC 438
Qy	338 ACCATGCTAGCTTGGCCATTCCTCAGGCGATCAGTTATGCCAAGCTGCCAACCTCCCT 397
Db	439 GCCATTGCCAGTCTCGCTATCCCTCAGGGAATCAGTTATGCCAAGCTAGCAAACTTACCA 498
Qy	398 CCAATCTTGGACTATATTTCGAGCTTTATACACCACTTGATTTATCCGATGATGGGTAGC 457
Db	499 CCCATTGTTGGTCTTTATTCAAGCTTCGTGGCCGCCATTATATCTCATGTTCTTGGAGT 558
Qy	458 TCGAGGATTTTGGCAGTGGGAGCTCTGGGGTGGATCGCTTCTGATGGGTTTCGATGTTG 517
Db	559 TCTAAGCATTTAGCAGTGGTCTCTCTCTATAGCATCACTTGTCTATGGGCTCAATGCTA 618
Qy	518 AGTAATGCCGTTGATCCCAATGAAGACCAAGCTTTTACCTCCACCTGGCTTTCACAGCT 577
Db	619 AGTGAAGTGTTCCTCCAACACAGACTCAATACTATATCTCAAACTGGCTTTTCACATCA 678
Qy	578 ACATTATTGCTGGTGTGTTTTCAGGCTGCTTGGGCTCTGTTAGGTTGGGTTGATCGG 637
Db	679 ACTTTCTTGTGGCCTCTTCCAAGCCTCCCTTGGCCTTCTCAGGCTGGGATTTGGGATT 738

Qy	638	GATTTTCTGTCAATGCAACCATAAATAGGGTTCATGGAGGAGCAGCAACGGTGGTGTGT	697
Db	739	GACTTTTGTCAAAAGGCAACATTTGGTTGGTTTCACTGCTGGTGCTGCAGTAATTTGTGTCA	798
Qy	698	CTGCAGCAACTAAAATCGATTCCTTGGCCTTGAGCATTTTCACCCATGAGCTGATATCATTA	757
Db	799	CTTCAACACAGCTTAAGGGTCTTCTTGGAAATCGTTCACTTCACCTGSCAAAATGCAATTCGTT	858
Qy	758	TCAGTGTAGCGCTCTGTGTTTTCAACCACAACTCATGAGTGGAGGTGGGAAAGTCTGTGTTA	817
Db	859	CCTGTCAATGCTCTGTGTATCAACACTAGATCCGAATGGTCATGGGAAACTATCTGTGATG	918
Qy	818	GGATGTGTTCATATTTCTTCCCTCTTAGCAACAAGATACTTCACGCAAAAAACGACCAAGG	877
Db	919	GGCCTTGGCTTCTTGATCATTTCTTAAACCAACAAGACACATTAGCATGAGGAAGCCAAAA	978
Qy	878	TTTTTTTTGGGTCTAGCAATGGCGCCATAGACGTCCGTTATATTGGGAAGTCTCTTGTTT	937
Db	979	CTTTTCTGGGATATCAGCTGCATCACCATTGGCATTCAGTTGTTATCTCAACTCTTCTTTGTA	1038
Qy	938	TATTTCTACTCACGCGCGAAGACACGGTGTGTGAAGTGATAGAGAACTGAAGAAGGGTTTG	997
Db	1039	TACGTCAATCAGAGCAAAAATCATGCCATCTCTTTTATTTAGGCCATCTGCCAAGGGTTTG	1098
Qy	998	AATCCACATCACTCAAAAATCTGGTATTTGTGTGGCCTTACATGACTACAGCTGTCAAA	1057
Db	1099	AATCCTCCTTCAGCGAATATGTTGTACTTTAGTGTCTCATCTCTCTTGCCATCAAG	1158
Qy	1058	ACTGGCAATGTCTGTGGCATCATATCACTTCGCGAAGGAATAGCAGTAGAAGAGCTTT	1117
Db	1159	ACTGGTATCATCACAGGGAATCTCTCCCTTACAGAAGGGATGCTGTAGGAGAACCTTT	1218
Qy	1118	GCAATGTATAAAAATTAACAATATTGATGGCAACAAAGAGATGATAGCTATTGGGACCATG	1177
Db	1219	GCATCTCTAAGAACTATCAAGTTAATGGGAAACAAAGAATATGTCGTATAGTTTTATG	1278
Qy	1178	AACGTAGTTGGTCTCTTTCACTCTTGCTACCTCAACACAGGACCAATTTCCGCTCGGCT	1237
Db	1279	AACATGCTGGTCTTGACCTCTTGCTATGTTTACAAACAGGGTCTTTCTCAAGATCTGCA	1338
Qy	1238	GTGAACTATAACGCTGGAATGCAAGACAGCAGCTTCGAAACATTAATGTCACTTGCAGTA	1297
Db	1339	GTAAACGTCAATGCTGGAGCGAAACAGCAGTTTCAAACATTTGTGATGGCATCAGCAGTT	1398
Qy	1298	ATGTTGACATTTGTTATTTCTTGACACCTGTGTTCCATTACACTCCCTCGTGGTGTCTATCA	1357
Db	1399	CTTGGGACCCCTTTGTTTTGTTGATGCCACTCTTCTATACACTCTTAATCTGATCTCAGCC	1458
Qy	1358	GCTATTAATCGTATCTGCAATCTCTGGACTCATAGATTATGAAGCAGCCATCCATCTATTT	1417
Db	1459	GCCATCATTTTAAACCGCTGTGATAGGCTCATTTGATTACCAACGGCTTACAAGCTTTGG	1518
Qy	1418	AAGTTGACAAATTTGACTTTTGGTGTGCATGATGTCATACATTTGGCGTGGTCTTTGGC	1477
Db	1519	AAAAGTTGACAAATTCGATTTTCTTCACTTGCAATGTGTTTCATTTCTTTGGCGTCTCTTTGTC	1578
Qy	1478	AGTCTTCAAAATTTGGCTTAGTCATAGCTATTGTAATATCTGTACTTCGGTACTCTCTATTT	1537
Db	1579	TCGTGCTCTTTGGCTTAGCAATAGCAGTGGAGTTTTCAAGTATCAAGATCTTTGTTGCAT	1638
Qy	1538	ATTGCAAGGCCAAGGACATTCGTTTTGGGCAACATTCGAAATTCGTGTATATACCGAAAT	1597
Db	1639	GTCAACAAGCCAAACACTTTTAGAGTTTCGGAACATCCCTGAGACTCAGATATATCAGAGT	1698
Qy	1598	GTTGAGCACTATCAAAATGCAAAACATGTTTCTCGGAATGCTAATTTCTAGAGATTGATGCA	1657
Db	1699	CTTAAAAAGATACAGAGAAGCCTCTAGAATCCCTGGTTTTCTTGATTTCTTTCAGTTGAATCT	1758
Qy	1658	CCAATTTACTTTGCTCAATGCCAGCTATTTAAGAGAAGGATCAACAGTTCGATTGATGA	1717
Db	1759	CCAATATACTTTGCCTCAATGCACTTACCTCAAGAAGGATCTCGAGATGCAACAGGGAA	1818
Qy	1718	GAAGAAGAAAGAAATTAAGCTACAGGGGAGACTAGTTTGCAGTATGTTATTAATTGATATG	1777

Db 1819 GAGGAAAATCGGATCAAGGAAAAACAATGAGAGAAACTTAAAGTGCAATTATCTTGACATG 1878  
 Qy 1778 AGTCTCTTGGAAACATTGATACAAAGTGGAAATAGATATGCTTGAAGAGGTGAAGAGATT 1837  
 Db 1879 ACAGCTGTGCTTCCATATAGACACAGATGGGATGGAATCGGTGTTGAACATTAGAGAGAGG 1938  
 Qy 1838 ACAGAGAGAGAGAGCTACAGCTGTGTTTGGTCAATCTCTGTAAGTGAAGTGAATGAAGAAA 1897  
 Db 1939 CTGAGAAATCAATCACTTCAGCTTGCTGCTGTAATCTCTGCGGAGTGTATGGAAG 1998  
 Qy 1998 CTGAACAAATCGAAGTTCCAAAATCATTTTAGGGAAGAAATGGATCTATCTGACTGTTGAA 1957  
 Db 1999 CTACACAAGTCCAGATCAATGACTCATTTGGGGCTGAGTGGGCTGATCGGGCTTTATCTAACAGTTGGT 2058  
 Qy 1958 GAGGCCGTT 1966  
 Db 2059 GAAGCTGTT 2067

RESULT 14  
 AB054645 Arabidopsis thaliana 2256 bp mRNA linear PLN 23-JAN-2001  
 LOCUS Arabidopsis thaliana mRNA for sulfate transporter Sultr3;4,  
 DEFINITION complete cds.  
 ACCESSION AB054645  
 VERSION AB054645.1 GI:12381948  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
 1 (sites)  
 Takanashi, H., Watanabe-Takahashi, A., Saito, K. and Yamaya, T.  
 cDNA for sulfate transporter Sultr3;4  
 Published Only in DataBase (2001)  
 2 (bases 1 to 2256)  
 Takanashi, H., Watanabe-Takahashi, A., Saito, K. and Yamaya, T.  
 Direct Submission  
 Submitted (19-JAN-2001) Hideki Takahashi, RIKEN Plant Science  
 Center; Hiroawa 2-1, Wako, Saitama 351-0198, Japan  
 (E-mail:hideki@postman.riken.go.jp, Tel:81-48-467-6899,  
 Fax:81-48-467-6857)

FEATURES  
 Location/Qualifiers  
 1..2256  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 95..2056  
 /codon\_start=1  
 /product="sulfate transporter Sultr3;4"  
 /protein\_id="BAB21264.1"  
 /db\_xref="GI:12381949"  
 /translation="MGHGTNRVEDMASPNNGTAGETVWEIHSVCLPDKKTAFOKLKRR  
 VGVFPPDQLQRFNTRNRNVLIGLSLFPFTWGSVDLKLRSDDVLSGTLIASL  
 AIPQGISYAKLANLPVIGLSFVFPPLIYAVLGSSEHFAVPSVLSIIVMGSMLES  
 VSTQSSILYKLKLAFTSTFPAGVFQASGLLRGLFMDLSKATLCGLFTAGAAVVS  
 QQLQGLGVHFGTKMQLVPMVSSVFNHRSEWSTFVMDIGLSILLTRHIVSMRKP  
 KLWISGAAPLASVITSTLLVLRISKTHAIPGHLKPLGNPPSLNMLYFSGAHIAL  
 AIKTGIITGLISTEGIVAGRTFASLKNQVNGNKEMMAIGFNMAGSCTSCVVTGDS  
 FRSASVYNAGAKATVSNIVMASAVLTLLFLMPLFYVTPNVITLAAIILTAVTGLDY  
 QAAKYLKWDKDPFFCLSPGVLVPLGLAIJAVSVIKILLHVTRPNWTSRGN  
 IPTQYISLGRYREASRIPGFLILAIESPYPANSTYLDQRLRWARBEENIKHKKI  
 GTLTKCIIIDMTAVSAIDTSGLEAVFELRRLKESQSLQLVLVNPVGTVMVKLHKSKI  
 EALGLSLYLTVGEAVADLISSTWKANGQP"

ORIGIN  
 Query Match 27.7%; Score 678.6; DB 8; Length 2256;  
 Best Local Similarity 60.5%; Pred. No. 1.5e-157;  
 Matches 1116; Conservative 0; Mismatches 729; Indels 0; Gaps 0;  
 Ov 123 TGACCAAGTCGAGGTTCCACCCGACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTGA 192

Db 171 TA C A G C G T T G T C T A C C C C G A A G A A A A C A C G G T T C A G A A A C T G A A A A A A C G C G T T G 230  
QY 183 A G G A G A C T T T C T T C C C T G A T G A C C C T T T G A G G C A G T T C A A G A C A A C G C A G C T T C C A A G A 242  
Db 231 G G A C G T G T T C T T C C C T G A T G A T C C G T T A C A G A G G T T A G G A A C C A A A C A T G G A G A A C A 290  
QY 243 A G T T C A T G C T T G C C T T C A G T T C T T C T T C C C A T T T T C G A A T T T T G C A A T T T T C C C A A T A C A C C T 302  
Db 291 G A G T G A T T C T A G G T C T T C A A G C T T G T T C C G A T A T T C A C A T G G G T T C T C A G T A C G A T C 350  
QY 303 T T C A G T T C T T G A A G C T G A C C T C A T A G C T G G C A T C A C C A T C G A T C G T A G C T T G G C C A T T C C T C 362  
Db 351 T C A A G C T T C T A G G T C C G A T G T T A T C T C T G G T C T C A C C A T T G C C A G T C G C C A T T C C T C 410  
QY 363 A G G G C A T C A G T T A T G C A A G C T C G C A A C C T C C C T C C A A T T C T T G A C T A T A T T C G A G C T 422  
Db 411 A G G A A T C A G T A T A G A A G C T A G C A A A C T T A C A C A T A T A G T T G G T C T T A T T C A A G C T 470  
QY 423 T T A T A C C A C C A T T A T T A T G C A T G A T G G G T A G C T C G A G G A T T T G G C A G T G G G A C T G 482  
Db 471 T T G T G C C A C C A C T G A T A T A T G C A G T T C T T G G A G T T C A A G A C A T C T T G C A G T T G T C C T G 530  
QY 483 T G C G G T T G A T C G C T T C T G A T G G G T T C G A T G T G A G T A T A T G C C G T T G A T C C C A A T G A A G 542  
Db 531 T C T C A A T A G C A T C A C T T G T A T G G G C T C A A T G C T A A G T G A A A G T T T C A C C A A C A C A A G 590  
QY 543 A C C C A A G C T T T A C C T C C A C C T G C T T T C A C A C T A C A T T A T T G C T G T G T T T T T C A G S 602  
Db 591 A T T C A A T T A T A T C T C A A A T T G C C T T C A C T T C A A C T T T C T T G C T G T G T C C C A A G 650  
QY 603 C T G C C T T G G G T C T G T T A G T T G G G T T G A T C T G T G A T T T C T G T C A C A T G C A A C C A A 662  
Db 651 C C T C C C T T G S C C T T C T A G S C T A G A T T A T A G A T T G A C T T T T G T C A A A G G C A A C T T T G A 710  
QY 663 T A G G T T C A T G G A G A G A C A G C A C C G T G T G T C T G C A G A A C T A A A A T C G A T T C T T G 722  
Db 711 T T G G T T C A C T G C T G T G T C A G T T A T C G T G T C A C T T C A A C A A C T A A A G G G T C T T C T T G 770  
QY 723 G C C T T C A G A C T T T C A C C A T G G A G C T A T A T A T A T A C A G T A T C A G T A T G C G T C T G T T T T C A C C 782  
Db 771 G A A T T G T T C A T T C A C T T G G C A A A T G C A A A T G T T C C T G T C A T G T C T T C T G T T T T C A A C C 830  
QY 783 A A A C T C A T A G T G G A G T G G A A G T C T G T G T T A G G A T G T G C T T C A T T T C T C C T C C 842  
Db 831 A T A G A T C T G A A G T C T T G S G A A C T A T T G T C A T G G G A T T G C T T C T T G A C A T T C T C T 890  
QY 843 T T A G C A C A G A T A C T T C A G C A A A A A C G A C C A A G G T T T T T T G G G T G C A G C A A T G G C G C 902  
Db 891 T A A C C A C A G A C A T T A G C A T G A G A A G C C A A A A C T G T T C T G G A T A T C T G C T G C A T C A C 950  
QY 903 C A T T G A C G T C C G T T A T A T T G G A A G T C T C T T G G T T A T T T C A C T C A C G C G A A G A C A G C 962  
Db 951 C T T T G G A T C G G T A T A T A T A C A C T C T G C T T G T A T A C C T C A T T A G A T C A A G A C T C A T G 1010  
QY 963 G T G T T G A A G T A T A G A G A C T C A A A A G G G T T T G A A T C C A C C A T C A C A A A C T C T G 1022  
Db 1011 C C A T C T T T C A T T G G A C A T C T A C C A A G G G T T G A T C C A C C T T C A T T A A A C A T G T G T 1070  
QY 1023 T A T T G T G C C T T A C A T G A C T A G C T G T C A A A A C T G G C A T T G C A T G T G T T G G C A T C A T A T 1082  
Db 1071 A C T T C A G T G T G T C A T C T T G C T C T T G C C A T C A A A C C G G A T A T C A T C A C C G G A T T C T T T 1130  
QY 1083 C A C T T C G G A A G A A T A G C A T A G A G A G C T T C C A A T G A T A A A A A T T A C A A T T G 1142  
Db 1131 C T C T T A C A A A G G A T T G C T G T A G G A G A A C C T T G C A T C T C T A A A G A A C T A T C A A G T T A 1190  
QY 1143 A T G C C A A A A G A G A T A T A G C T A T T T G G A C C A T G A A C G T A G T T G T T C T T T C A C C T C T T 1202  
Db 1191 A T G A A A C A A A A A T G A T G C T A T A G G T T T A T A A C A T G G C T G T T C T T G C A C C T C T T 1250  
QY 1203 G C T A C C T C A A A C A G A C C A A T T T T C G G T T C G G C T G T G A A C T A T A A C G C T G G A T G C A A G A 1262

Db 1251 G C T A T G T T A C A A C A G G G T C T T T C T C T C G A T C T C G G T A A A C T A T A A T G C T G A G C G A A A A 1310  
QY 1263 C A G C A G C T T C C A A C A T A A T A T A T G T C A C T T G C A G T A A T G T T G A C A T T G T T A T T C C T G A C A C 1322  
Db 1311 C A G C G G T T C T A A C A T T G T A T G G C C T C T G C A G T T C T A G T A C C C T C T G T T T T T G A T G C 1370  
QY 1323 C C T T G T C A T T A C A C T C C C C T G G T G G T C T A T A G C T A T T A T C G T A T T A T C G T A T C T G C A A T G C T T G 1382  
Db 1371 C A C T C T C T A T T A C A C T C C T A A T G T G A C C T A G C T G C T A T C A T C T T A A C C G G G T G A T A G 1430  
QY 1383 G A C T C A T A G A T T A T A A G C A G C C A C C A T C T A T T A A G G T T G A C A A A T T T G A C T T T G T G G 1442  
Db 1431 G C C T C A T T A T A T C A A G C G C T T A A A C T T T G A A A G T T G A C A A T T C G A A T T C T T C A 1490  
QY 1443 T G T C A T A G T C A T A C A T T G C G T G G T T T T G C A G T G T T G A A A T T G G C T T A G T C A T A G 1502  
Db 1491 C T T G C T T G T T C C T C T T T G C G T T C T T T G T T C T T G T A C C T C T T G C C T A G C A A T A G 1550  
QY 1503 C T A T T G T A A T A T C T G T A C T T C G G G T A C T T C T A T T A T T G C A A G G C C A A G G A C A T T C G T T T 1562  
Db 1551 C A G T G C A G T T T C A G T T A T C A A G A T C T T G T G C A G T A A C C A G G C C A A C A C T T C A G A A T 1610  
QY 1563 T G G G C A A C A T T C C A A A T T C T G T G A T A T A C C G A A A T G T T G A G C A C T A T C A A A T G C A A A C 1622  
Db 1611 T T G G A A T A T C C C A G G A C T C A G A T T A T C A G A G T C T T G G A A G A T A C A G A A G C T T C A A 1670  
QY 1623 A T G T T C C T G G A A T G C T A A T T C T A G A G A T T A T G C A C C A A T T T A C T T T G C C A A T G C A G C T 1682  
Db 1671 G A A T C C C A G G C T T C C A A T C C T T G C T A T T G A T C C T A T A T T T C G C C A T A G C A C T T 1730  
QY 1683 A T T A A G A A A A G A A T C A A A G T G G A T T G A T G A A G A A G A A A A A T T A A A G C T A C A G 1742  
Db 1731 A T C T T C A A G A C A G A A T T T G A T G G G C C A G G A A G A G A A A A T C G G A T A A A G A G A A C A 1790  
QY 1743 G G A C A C T A G T T T G C A G T A T G T T A A T T G A T A T A G T A T A G T G C T G T T G G A A C A T T G A T A C A A 1802  
Db 1791 A T G G T A C T A C C T T G A A A T G C A T A T T C T T G A C A T G A C A G C T G T G T C G G A T A G A C A A A 1850  
QY 1803 G T G G A A T A A G T A T G C T T G A A G A G T T G A A G A A T T A C A G A G A A G A G A G C T A C A G C T T G 1862  
Db 1851 G C G G C T T G A A C G G T G T T T G A C T T A G A G A G A C T C G A A G A C A A T C T C T T C A G C T T G 1910  
QY 1863 T T T T G C T C A A T C C T G T A A G T A A G T G A G A A A A C T G A A C A A A T C G A A A A T C G A A G T T C C A A A A T C 1922  
Db 1911 T G C T G T G A A T C C T G T G G G A C T G T G A T G G A A A A G C T A C A A A G T C C A A A G T C A T T G A G G 1970  
QY 1923 A T T A G G A A G A A A T G G A T C T A T C T G A C T G T T T G A A G A G C C G T T G 1967  
Db 1971 C A T T G G G T C T G A T G G A C T T T A T C T A A C A G T T G G T G A A G C T G T G 2015

RESULT 15  
AF355602  
LOCUS  
DEFINITION  
AF355602  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AF355602 2640 bp mRNA linear PLN 13-JUL-2004  
Zea mays sulfate transporter ST1 mRNA, complete cds.  
AF355602.1 GI:13625940  
Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 2640)  
Hopkins, L., Parmar, S., Bouranis, D.L., Howarth, J.R. and  
Hawkesford, M.J.  
Coordinated Expression of Sulfate Uptake and Components of the  
Sulfate Assimilatory Pathway in Maize  
Plant Biol. 6 (4), 408-414 (2004)  
2 (bases 1 to 2640)  
Hopkins, L., Bouranis, D.L. and Hawkesford, M.J.  
Full length cDNA of a high-affinity type sulfate transporter from  
Zea mays

JOURNAL REFERENCE	Unpublished
3 (bases 1 to 2640)	Hopkins, L., Bouranis, D. L. and Hawkesford, M. J.
AUTHORS	Direct Submission
TITLE	Submitted (01-MAR-2001) Agriculture and Environment Division,
JOURNAL	IACR-Rothamsted, Harpenden, Hertfordshire AL5 2JU, UK
FEATURES	Location/Qualifiers
source	1..2640
	/organism="Zea mays"
	/mol_type="mRNA"
	/db_xref="taxon:4577"
	120..2096
	/codon_start=1
	/product="sulfate transporter ST1"
	/protein_id="AAK35215.1"
	/db_xref="GI:13625941"
	/translation="MPERTVSDGDDNGTDVASRTSSHRETEHGKVGKGVPPKKNLL AETSDAVKETPPADDPRLQYKDQPSKKIWLGLQHPVLDNSRHVSLGKFKGDFATG TLTASUCIPDITGLEYRLAUPAEVGLSYSPFLIYVAGSSRSDIAIGPVAVVSLLLG TLQNEEDPOTTEFLVKRNPATFAFGVQAALGFRLGFIIEFISHAAIVGFMAGA AIIATLQQLGFGIAGIATPKKSDIVSMKSVGNVHGNWQITLIGATFLAPLLVAF YIGKRNKLFWSAIAPLTVSIITFVYITRADKIGVAIVKNIRKGINPPSPASLIIVF TGPLYATGFKIGIVAGMIGTEAIAIGRTPAALKDYRIDGNKEMWALGTNIVGSLTS CVYATGFSRSSAVNWKDGLTAVSNVMSIVNMLTLLLTLPFLKTPFNAILSSIIILSA VLGLDYESAYLWPKAGDPLFACMGAFGVIFSSVEYGLLIAVIALSLAKILLQVTRP EQQODOKLTTEFLIVDLSPVIDIDTSGIHABELAKALEKRKIQLVLTNPGPAVIOK LRSAKFTMDIMEDNIFLITVGDAVKKPAKPNVDV"
ORIGIN	
Query Match	27.7%; Score 677.4; DB 8; Length 2640;
Best Local Similarity	60.8%; Pred. No. 3e-157;
Matches 1123; Conservative	0; Mismatches 721; Indels 3; Gaps 1;
Qy	122 GTGCACCAAGTCGAGGTTCCACGCCACACGCCGTTTTTCAAGTCTCTAAAGTACTCTTTG 181
Db	213 GGGTACAAAGTCGCGCGTCCACCGGAGAAACCTCTCTCGCCGAGATCTCTGACGGGTG 272
Qy	182 AAGGAGACTTTCTTCCTCGATGACCCCTTCGAGGCAGTTCAAGAACAGCAGCTTCCAAG 241
Db	273 AAGGAGACGTTCTTCGCCGACGACCCGCTGGCGCAGTCAAGGACAGCCAGGTCGAG 332
Qy	242 AAGTTCATGCTTG3CCTTCAGTTCTTCTTCCCAATTTTCGAATGGGCTCCCAATACACC 301
Db	333 AAGATCTGGCTCG3CCTCGACGACATCTTCCCGTGTCTGACTGGACAGGCATTACTCC 392
Qy	302 TTTTCAGTTCTTGAAGCTGACCTCATAGCTGGGCATCAACATCGGTAGCTTGGCCATTCCT 361
Db	393 CTGGCAAGTTCAGGGTGACTTCATCGCTGGGCTCAACATTCGACGCTCTGCATACCC 452
Qy	362 CAGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAATTTCTTGAGACTATATTCGAGC 421
Db	453 CAGACATCGGTATTCTAAGTTGTCTTACTTCCAGCAGAGGTGGGACTATACAGTAGC 512
Qy	422 TTTTATACCACTTATGATTATGCGATGATGGGTAGCTTCGAGGATTTTGGCAGTGGGGACT 481
Db	513 TTCGTGCGCTCTGTATATACGCAGTGAATGGCGAGCTCCAGGACATAGCCATCGGCCCA 572
Qy	482 GTGGCGGTTCGATCGCTTCTGATGGGTTCCATGTTGAGTAATGCCGTTGATCCCAATGAA 541
Db	573 GTGGCCGTGTGTGCTGTGCTGCTCGGTACTCTCCCTTCAGAATGAGATCGACCCCAAGACA 632
Qy	542 GACCCAAAGCTTTTACCTTCCACTGGCTTTTCACAGCTACATTTATTCGTGGTGTGTTTTTCAG 601
Db	633 CACCCGCTCGAGTACAGCGCGCTAGCCCTTCACAGCAACCTTTTTTCGACGGGGTCACTTCAG 692
Qy	602 GCTGCTTTGGGTCTGTTTTAGGTGTGGGTTTGATCGTGGATTTTCTGTGTCATGCAATCCATA 661
Db	693 GCAGCCCTCGGATCTTCAGGCTAGGGTTTATCATAGAGTTCTTGTCTCATGTGCGCAATT 752
Qy	662 ATAGGGTTCAATGGAGAGAGCAGCCACGGTGGTGTGTCGAGCAACTTAAATCGATTCCT 721
Db	753 GTCGGATTCATGGCCGGCTGCGCATCACCAATTCGCTTCAGCAGCTGGAAGGATTCCTT 812

Qy	1799	ACAAAGTGGAAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAG	1858
Db	1893	ACAAAGTGGAAATCCATGCTTTGGAGGAGTTGGCGAAAGCTCTTGAAAAACGCAAAATTTCAG	1952
Qy	1859	CTTGTTTGGTCAATCCTGTAAGTGAAGTGAAGAAACTGAACAAATCGAAGTTCCAA	1918
Db	1953	CTGGTTCTTACCAATCCCGGCGCGGGTGATCCAGAAGCTCCGCTCAGCGAAATTTCAG	2012
Qy	1919	AATCAATTTAGGGAAGAAATGGATCTATCTGACTGTTGAAGAGGCCGT	1965
Db	2013	GACATGATTGGTGAAGACAACATATTCCTCAGGTCGGCGACGCTGT	2059

Search completed: August 31, 2005, 06:33:50  
Job time : 10593 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 19:18:05 ; Search time 1312 Seconds

(without alignments)  
11049.883 Million cell updates/sec

Title: US-10-762-049-17

Perfect score: 2449

Sequence: 1 gcacgagctagctgcacat.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001as:\*

5: Geneseq2001bs:\*

6: Geneseq2002as:\*

7: Geneseq2002bs:\*

8: Geneseq2003as:\*

9: Geneseq2003bs:\*

10: Geneseq2003cs:\*

11: Geneseq2003ds:\*

12: Geneseq2004as:\*

13: Geneseq2004bs:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2449	100.0	2449	3 AAZ50489	Aaz50489 Soybean s
2	1041.2	42.5	1977	3 AAC43310	Aac43310 Arabidops
3	1038	42.4	2112	3 AAC43093	Aac43093 Arabidops
4	1018.2	41.6	1977	6 ABZ13697	Abz13697 Arabidops
5	1018.2	41.6	1977	8 ADA68514	Ada68514 Arabidops
6	846	34.5	1974	8 ADA70002	Ada70002 Rice gene
7	765.2	31.2	1981	3 AAZ50482	Aaz50482 Corn sul
8	720.6	29.4	2067	3 AAZ50488	Aaz50488 Soybean s
9	662.4	27.0	2006	3 AAC42229	Aac42229 Arabidops
10	660.2	27.0	1896	8 ADA68554	Ada68554 Arabidops
11	655.4	26.8	2279	3 AAZ50481	Aaz50481 Corn sul
12	650.4	26.6	1989	8 ADA70466	Ada70466 Rice gene
13	647.2	26.4	1947	8 ADA69770	Ada69770 Rice gene
14	640.4	26.1	2311	3 AAZ50490	Aaz50490 Wheat sul
15	634.4	25.9	1971	6 ABQ82719	Abq82719 Arabidops
16	630.4	25.7	1940	8 ADA71315	Ada71315 Rice gene
17	600.2	24.5	2000	8 ADA70110	Ada70110 Rice gene
18	559.8	22.9	1896	8 ADA69681	Ada69681 Rice gene
19	455.2	18.6	1983	8 ADA69804	Ada69804 Rice gene
20	401.4	16.4	1980	12 ADJ40066	Adj40066 Plant cDN

21	398.8	16.3	1974	8 ADA70351	Ada70351 Rice gene
22	324	13.2	780	3 AAZ50484	Aaz50484 Arabidops
23	217.2	8.9	2058	6 ABZ12912	Abz12912 Arabidops
24	217.2	8.9	2058	8 ADA68540	Ada68540 Arabidops
25	216.2	8.8	1240	3 AAZ50483	Aaz50483 Corn sul
26	210.8	8.6	500	10 ABX61505	Abx61505 Arabidops
27	208.8	8.5	490	10 AD881920	Ad881920 Arabidops
28	205.6	8.4	2022	3 AAZ50491	Aaz50491 Wheat sul
29	184.8	7.5	596	12 ADJ43778	Adj43778 Plant cDN
30	148.8	6.1	552	12 ADJ44485	Adj44485 Plant cDN
31	129.6	5.3	699	9 ACL17188	ACL17188 DNA clone
32	124.6	5.1	695	9 ACL17186	ACL17186 DNA clone
33	123.2	5.0	680	9 ACL17191	ACL17191 DNA clone
34	118	4.8	493	3 AAZ50487	Aaz50487 Rice sul
35	117.4	4.8	484	3 AAZ50485	Aaz50485 Rice sul
36	115.2	4.7	658	9 ACL17184	ACL17184 DNA clone
37	104.4	4.3	630	9 ACL17187	ACL17187 DNA clone
38	100.2	4.1	264	6 ABL74697	Ab174697 Corn tass
39	97.4	4.0	2241	3 AAZ29466	Aaz29466 P. chryso
40	94.8	3.9	594	9 ACL17185	ACL17185 DNA clone
41	93.4	3.8	581	9 ACL17189	ACL17189 DNA clone
42	90.4	3.7	2441	11 ADM57788	Adm57788 Murine pr
43	90.4	3.7	2441	12 ADE86099	Ade86099 cDNA enco
44	89.4	3.7	2459	6 ABQ59208	Abq59208 Listeria
45	89.4	3.7	2632	6 ABQ70888	Abq70888 Listeria

ALIGNMENTS

RESULT 1

AAZ50489

ID AAZ50489 standard; cDNA; 2449 BP.

XX

AC AAZ50489;

XX

DT 23-MAY-2000 (first entry)

XX

DE Soybean sulphate permease-2 cDNA clone.

XX

KW Sulphate Permease; sulphate assimilation protein; soybean; probe;

KW mapping; marker; plant breeding; chimeric gene; transgenic plant;

KW antibody; screen; ss.

XX

OS Glycine max.

XX

FH Key

FT CDS

FT 2..2044

FT /\*tag= a

FT /product= "Soybean sulphate permease-2"

FT /note= "Derived from clone sfl1.pk0043.g10"

XX

PN WC200004154-A2.

XX

PD 27-JAN-2000.

XX

PF 13-JUL-1999; 99WO-US015810.

XX

PR 14-JUL-1998; 98US-0092833P.

XX

PA (DUPO ) DU PONT DE NEMOURS & CO E I.

XX

PI Allen SM, Falco SC, Thorpe CJ;

XX

DR WPI; 2000-195025/17.

DR P-PSDB; AAY44943.

XX

PT Nucleic acid fragments encoding sulfate assimilation proteins in plants

PT and seeds useful as probes for isolating cDNAs and genes encoding

PT homologous proteins, in producing transgenic plants.

XX

PS Claim 3; Page 52-53; 79pp; English.

XX



CC The present sequence is a cDNA clone encoding soybean sulphate permease,  
CC a sulphate assimilation protein. This sequence is obtained from sf11  
CC library, clone sf11.pk0043.g10, derived from soybean immature flower.  
CC This sequence is used as a probe to isolate other plant sulphate  
CC assimilation proteins, for genetic and physical mapping of related genes  
CC and as markers of traits linked to the gene. This is useful for plant  
CC breeding and to construct chimeric genes, used to create transgenic  
CC plants with altered levels of sulphate permease. The sulphate permease  
CC peptides are useful for producing antibodies, that are used to screen and  
CC isolate cDNA clones  
XX  
SQ

Sequence 2449 BP; 719 A; 470 C; 543 G; 717 T; 0 U; 0 Other;

Query Match 100.0%; Score 2449; DB 3; Length 2449;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GCACGAGCTAGCTCGCACATTAAGTTATATATAACACATATTGCTTGCTTAGAAATACTAT	60
Db	1	GCACGAGCTAGCTCGCACATTAAGTTATATATAACACATATTGCTTGCTTAGAAATACTAT	60
Qy	61	TATTGAAGATATGGGAGTGTAGATTATAGTACCCTTTGGGCATGAACAACTTCGAGAG	120
Db	61	TATTGAAGATATGGGAGTGTAGATTATAGTACCCTTTGGGCATGAACAACTTCGAGAG	120
Qy	121	AGTGCACCAAGTCGAGGTTCCACCGCCACACGCGTTTTTCACAGTCTCTAAAGTACTCTTT	180
Db	121	AGTGCACCAAGTCGAGGTTCCACCGCCACACGCGTTTTTCACAGTCTCTAAAGTACTCTTT	180
Qy	181	GAAGGAGACTTTCTTCCTGATGACCCCTTTGAGGCGAGTTCAGAAACAGCAGCTTCCAA	240
Db	181	GAAGGAGACTTTCTTCCTGATGACCCCTTTGAGGCGAGTTCAGAAACAGCAGCTTCCAA	240
Qy	241	GAAGTTCACTCTGGCCTTCAGTTCCTTCCCTCCCATTTTCGAATGGCTCCCAAAATACAC	300
Db	241	GAAGTTCACTCTGGCCTTCAGTTCCTTCCCTCCCATTTTCGAATGGCTCCCAAAATACAC	300
Qy	301	CTTTTCAGTCTTGAAGAGCTGACCTCATAGCTGCGCATCACCATCGCTAGCTTGGCCATTCC	360
Db	301	CTTTTCAGTCTTGAAGAGCTGACCTCATAGCTGCGCATCACCATCGCTAGCTTGGCCATTCC	360
Qy	361	TCAGGGGATCAGTTATGCCAAGCTCGCCAACTCCCTCCCAATTCCTTGGAGCATATATTCGAG	420
Db	361	TCAGGGGATCAGTTATGCCAAGCTCGCCAACTCCCTCCCAATTCCTTGGAGCATATATTCGAG	420
Qy	421	CTTTATACCACTTGAATTTATCGGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAC	480
Db	421	CTTTATACCACTTGAATTTATCGGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAC	480
Qy	481	TGTGGCGTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAAATGCCGTTGATCCCAATGA	540
Db	481	TGTGGCGTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAAATGCCGTTGATCCCAATGA	540
Qy	541	AGACCCAAAGCTTTACCTCACCTGGCTTTCACAGCTACATTAATTTGCTGGTGTGTTTCA	600
Db	541	AGACCCAAAGCTTTACCTCACCTGGCTTTCACAGCTACATTAATTTGCTGGTGTGTTTCA	600
Qy	601	GGCTGCTTGGGTCTGTTTATAGTTGGGTTGATCGTGGATTTTCTGTCACATGCAACCAT	660
Db	601	GGCTGCTTGGGTCTGTTTATAGTTGGGTTGATCGTGGATTTTCTGTCACATGCAACCAT	660
Qy	661	AATAGGGTTTCATGGGAGGAGCAGCCAGGTTGGTGTCTGCGACAACTAAAATCGATTCT	720
Db	661	AATAGGGTTTCATGGGAGGAGCAGCCAGGTTGGTGTCTGCGACAACTAAAATCGATTCT	720
Qy	721	TGGCCTTGAGCATTTACCCATGGAGCTGATATCATATCATGATGCGGCTCTGTTTTAC	780
Db	721	TGGCCTTGAGCATTTACCCATGGAGCTGATATCATATCATGATGCGGCTCTGTTTTAC	780
Qy	781	CCAAACTCATGAGTGGAGGTGGGAAAGTGTCTGTTAGGATGTTCTTCAATTTCTTCTCT	840
Db	781	CCAAACTCATGAGTGGAGGTGGGAAAGTGTCTGTTAGGATGTTCTTCAATTTCTTCTCT	840

Qy	841	CCTTAGCAACAGATACTTCTAGCAAAAAACGACCAAGTTTTTTTTGGGTGTCTAGCAATGGC	900
Db	841	CCTTAGCAACAGATACTTCTAGCAAAAAACGACCAAGTTTTTTTTGGGTGTCTAGCAATGGC	900
Qy	901	GCATTTGAGTGGTATATTGGGAAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGAACA	960
Db	901	GCATTTGAGTGGTATATTGGGAAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGAACA	960
Qy	961	CGGTGTTGAAGTGTAGGAGAACTGAAGAAGGGTTTGAATCCACCATCACTCACAATCT	1020
Db	961	CGGTGTTGAAGTGTAGGAGAACTGAAGAAGGGTTTGAATCCACCATCACTCACAATCT	1020
Qy	1021	GGTATTTGTCGCTTACATCACTACAGCTGTCAAACTGGCATTTGCTTGGCATCAT	1080
Db	1021	GGTATTTGTCGCTTACATCACTACAGCTGTCAAACTGGCATTTGCTTGGCATCAT	1080
Qy	1081	ATCAGCTTGGGAAGGAATAGCAGTAGGAGAACTTTCGAATGTATAAAAAATTACAATAT	1140
Db	1081	ATCAGCTTGGGAAGGAATAGCAGTAGGAGAACTTTCGAATGTATAAAAAATTACAATAT	1140
Qy	1141	TGATGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTACCTTC	1200
Db	1141	TGATGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTACCTTC	1200
Qy	1201	TTGCTACCTCAACACAGGACCAATTTTCGCGTTTCGCTGTGAATCTATAACCGTGGATGCAA	1260
Db	1201	TTGCTACCTCAACACAGGACCAATTTTCGCGTTTCGCTGTGAATCTATAACCGTGGATGCAA	1260
Qy	1261	GACAGAGCTTCCAAATTAATTAATGTCACCTTCGAGTAATGTTGACATTTGTTTCTGTAC	1320
Db	1261	GACAGAGCTTCCAAATTAATTAATGTCACCTTCGAGTAATGTTGACATTTGTTTCTGTAC	1320
Qy	1321	ACCTTGTTCATTAACACTCCCTGGTGTCTATCAGCTATTATCGTATCTGCAATGCT	1380
Db	1321	ACCTTGTTCATTAACACTCCCTGGTGTCTATCAGCTATTATCGTATCTGCAATGCT	1380
Qy	1381	TGGACTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTGT	1440
Db	1381	TGGACTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTGT	1440
Qy	1441	GGTGTGATGAGTGATACATTTGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAT	1500
Db	1441	GGTGTGATGAGTGATACATTTGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAT	1500
Qy	1501	AGCTATTGTAATCTGTACTTTCGCTGCTCTATTTATTGCAAGGCCAAGGACATTCCT	1560
Db	1501	AGCTATTGTAATCTGTACTTTCGCTGCTCTATTTATTGCAAGGCCAAGGACATTCCT	1560
Qy	1561	TTTGGGCAACATTCGCAAAATCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAA	1620
Db	1561	TTTGGGCAACATTCGCAAAATCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAA	1620
Qy	1621	ACATGTTCTCGGAATGCTAAATCTAGAGATTGATGCACCAATTTACTTTTGGCAATGCCAG	1680
Db	1621	ACATGTTCTCGGAATGCTAAATCTAGAGATTGATGCACCAATTTACTTTTGGCAATGCCAG	1680
Qy	1681	CTATTTAAGAGAAAGGATCAACAGGTGGAATGATGAAAGAAAGAAAGAAATTTAAAGCTAC	1740
Db	1681	CTATTTAAGAGAAAGGATCAACAGGTGGAATGATGAAAGAAAGAAAGAAATTTAAAGCTAC	1740
Qy	1741	AGGGGAGACTAGTTTGCAGTATGTTTATATGATATGATGCTGTGGGAAACATTTGATAC	1800
Db	1741	AGGGGAGACTAGTTTGCAGTATGTTTATATGATATGATGCTGTGGGAAACATTTGATAC	1800
Qy	1801	AACTGGGAATGATGCTTGAAGAGGTGAAGAGATTACAGAGAGAGAGAGCTACAGCT	1860
Db	1801	AACTGGGAATGATGCTTGAAGAGGTGAAGAGATTACAGAGAGAGAGAGCTACAGCT	1860
Qy	1861	TGTTTTGGTCAATCTCTGTAAGTGAAGGAACTGAAACAAATCGAAGTTTCCAAA	1920
Db	1861	TGTTTTGGTCAATCTCTGTAAGTGAAGGAACTGAAACAAATCGAAGTTTCCAAA	1920
Qy	1921	TCATTTAGGGAAGAAATGGATCTATCTGACTGTTGAAAGAGGCCGTTGGAGCATGCAACTT	1980





PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	08-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	04-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	22-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 42.5%; Score 1041.2; DB 3; Length 1977;		
Best Local Similarity 72.4%; Pred.No.2.1e-246;		
Matches 1365; Conservative 0; Mismatches 518; Indels 3; Gaps 1;		
QY	125	CACCAAGTCGAGGTTCCACCGCACACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG 184
DB	58	CACACGGTGGAGGCTCCACAACCTCAACCGTTCTTTGAAGTCACTTCAGTACTCAGTGAAG 117
QY	185	GAGACTTTCTCCCTGATGACCGTTTGAGGCGAGTTCAAGAACAGCCAGCTTCCAAAGAAG 244
DB	118	GAACCTCTGTTTCCAGACGACCCCTTTTACACAATTTTAAAGAACCAAAATGCATCAAGAAA 177
QY	245	TTTCATGCTTTGGCCTTCAGTTCTTTCCCAATTTTTCGAATGGCTCCCAAAATACACCTTT 304
DB	178	TTTGTGTTAGGCTCAATATCTTCTCCCGATTTTCGAATGGGACCAAGCTACAAATCTC 237
QY	305	CAGTTCTTGAAGCTGACCTCATAGTGGCATCACCATCGTAGTCTAGCTTGGCCATTCCTCAG 364
DB	238	NAAGTTCTTCAATCAGATCTCATCGCGGAATCACCATCGTAGTCTGCCATCCCTCAG 297
QY	365	GGCATCAGTTATGCCAAGCTCGCGAACCTCCCTCCAAATTCCTGGACTATATTCGAGCTTT 424
DB	298	GGCATCAGTTACGCCAAACTTGTAACTTTGCCCCCAATTTCTTGGCCCTTTATTCGAGTTTT 357
QY	425	ATACCACCAATTCATTATTCGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACTGTG 484
DB	358	GTACCGCCATTGGTATACCGGTGCTAGGGAGTTCAAGGAGCTTAGCGGTGGGAACGGTT 417
QY	485	CGCGTTGGATCGCTTCTGATGGGTTTCGATGTTGAGTAATGCCGTTGATCCCAATGAAGAC 544
DB	418	CGCGTTGGCTCTCTGTTGACAGTGCATGCTGAGCAAGAAGTTGATGCTGAGAAAGAT 477
QY	545	CCAAAGCTTTACTTCACCTGCTTTCACAGCTACATTAATTTGCTGGTGTGTTTTTCAAGCT 604
DB	478	CCTAAGCTTTTACTTTTCACTTTCACCGCACCTTTTTTCCCGCGCGTTTCTCGAAGCC 537
QY	605	GCCTTGGGTCTCTTTAGGTTGGGTTGATCGTGGATTTTCTGTCACTGCAACCAACCAATA 664
DB	538	TCTCTGGAAATTTTTCAGGTTAGGGTTTCATAGTGGACTTCTATCGCATGCCAAGATAGTA 597
QY	665	GGGTTTCATGGGAGGAGCAGCCACGGTGGTGTGTCTGTCAGCAAACTAAAAATCGATTTTGGC 724
DB	598	GGATTCATGGGAGGAGCAGCGAGCGGTGGTGTGAGTCTGCAACAGCTTTAAGGTTATTTTCGGA 657
QY	725	CTTGAGCATTTTCAACCATGGAGCTGATATCATATCAGTGATGGCTCTGTTTTTCAACCAA 784
DB	658	CTTAAACATTTTCAACAGACTCTACCGATGTTATCTCTGTGTCATCGGTTTCGTTTCTCCCAA 717
QY	785	ACTCATGAGTGGAGGTGGGAAAGTGTGTGTTAGGATGTGTCTTCAATTTCTTCTCTCTT 844
DB	718	ACTCAGAGTGAGATGGGAAAGTGGCGTTCTTGATGTGGTTTTCTATTTCTTCTCTCTC 777
QY	845	AGCACAAGATACTTCAGCAAAAAACGACCAAGGTTTTTTTTTGGGTGTTCAGCAATGGGCCA 904

Db 778 TCACACAGATATTTACAGATCAAGAAACCAAAATCTTTGGGTGGCGGATGGCTCCT 837  
Qy 905 TTGACGTCGGTTATTTGGGAAGTCTCTTGCTTTATTTCTACACGCCGAGAAACGACGGT 964  
Db 838 TTGACCTCAGTGATTTCTGGAAAGTCTCTGGTTTACTTCACTCAGCGTGAGAGACATGGT 897  
Qy 965 GTTGAAGTGATAGGAAGCACTGAAGAGGGTTTGAATCCACATCACTCAAAATCTGGTA 1024  
Db 898 GTTCAAGTGATAGGGACCTGAAGAAAGGGTTGAATCCACTCTCCGGTTCTGTATCTATC 957  
Qy 1025 TTTGTGTCGCTTACATGACTTACAGCTGTCAAACTGGCAATGTCTGGCATCATATCA 1084  
Db 958 TTTACTTCCCTTACATGTCCACAGCTGTCAAACTGGCCTCATCTGCAATCATTTGCT 1017  
Qy 1085 CTTGCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAATTTACAATTTGAT 1144  
Db 1018 CTCGCTGAAGAGTAGCAGTGGGAGGAGTTTTCGATGTTTCAAGAACTACAACATAGAC 1077  
Qy 1145 GGCACAAAGAGATGATAGTATTTGGGACCATGAACGTAGTTGGTTCTTTTCACTCTTGC 1204  
Db 1078 GGGAAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCTTTTCACTCTGT 1137  
Qy 1205 TACCTCAACACAGACCAATTTTCGCTTCGCTGTGAACCTATAACGCTGGATGCAAGACA 1264  
Db 1138 TACCTCAACACCGACCAATTTTCAAGGTTCGCGAGTGAACCTACACGGGGTTGCAAGACC 1197  
Qy 1265 GCAGCTTCCAAACATTAATATGTCATCTTGCAGTAAATGTTGACATTTTATTCCTGCACACC 1324  
Db 1198 GCAATGTCCAAACATAGTATGATGGCGATTTCGGTTATGTTTCACACTCTCTTCTCACACCG 1257  
Qy 1325 TTGTTCCATTTACACTCCCTCGTGGTGTCTATCAGCTATTTATCGTATTCGAATGCTTTGA 1384  
Db 1258 CTTTTTCACTACACACCACTCGTCTCTCTGCCATCATATATCGCAATGCTCGGA 1317  
Qy 1385 CTCATAGATTATGAAGCAGACCATCCATCTATTTAAAGGTTGCAAAATTTGACTTTTGTGTG 1444  
Db 1318 CTATTGACTATCAGCTGCGATCCATCTCTGAAAGTTGCAAGTTTCGACTTTCCTCGTC 1377  
Qy 1445 TGCATGAGTGCATACATTTGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTTAGTCATAGCT 1504  
Db 1378 TGCATGAGCGCTACGTTCGGTGGTATTCGCGAGTGTAGATTTGCACTCGTGTAGCG 1437  
Qy 1505 ATTGTAATATCTGACTTCCTGGTACTTCTATTTATTCGAAGGCGCAGGACATTCGTTTTG 1564  
Db 1438 GTGCGGATATCTATAGCGAGGTTGTTGCTGTTGTGCGAGGCCAAAACTGCGGTTGAAG 1497  
Qy 1565 GGCACCAATTCCAAATTCGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACAT 1624  
Db 1498 GGAACATACCAAACAGCATGATCTATAGGAACACTGAGCAGTACCCATCATCAAGACC 1557  
Qy 1625 GTTCTCTGGAAATGCTAATTTCTAGAGATTGATGCACCAATTTTACTTTGCAATGCCAGCTAT 1684  
Db 1558 GTTCTCTGGTATCTCATCTTGGAGATTGATGCTCCCATCTCTTTGCTAATGCCAGTTAC 1617  
Qy 1685 TTAAGAGAAAGGATCAAGAGTGGAATTGATGAAGAGAAAGAAATTTAAAGCTACAGGG 1744  
Db 1618 TTGCGTGAAGAAATCAAGAGTGGAATTGATGAAGAGAGAGAGAGAGTTAAACAAATCAGGA 1677  
Qy 1745 GAGACTAGTTTGCAGTATGTTAATTTGATATGAGTGTCTGTTGGAACATTTGATACAAGT 1804  
Db 1678 GAGAGCAGCTTACAAATATATATATCTGATGTCGATGTCGAGCTGTTGGTAAATTCGACACAAGC 1737  
Qy 1805 GGAATAGTATGCTTGAAGAGGTGAAGAAATTTACAGAGAGAAGAGAGCTTACAGCTTGTT 1864  
Db 1738 GGTATTAGCATGATGGTGGAAATTAAGAAAGTCAATTGACAGGAGAGCGTTAAAGTTGGTA 1797  
Qy 1865 TTGCTCAATCTGTAAGTGAAGTGAAGAAACTGAACAAATCGAAGT---TCCAAAT 1921  
Db 1798 TTGTCAAATCCAAAGAGAGAGGTCGTGAAGAAATTAACACAGATCCAAATTCATCGGTGAT 1857  
Qy 1922 CATTTAGGGAAGAAATGGAATCTATCTGACTGTTGAGAGGCCGCTTGGAGCATCAACTTC 1981

Db 1858 CATTTGGCAAGAGTGGATGTTCTTAACGGTAGGAGAAAGCAGTGGAGGCTTGTAGCTAC 1917  
Qy 1982 AATCTACGTGCAGCAAGCAAAACGAAACC 2007  
Db 1918 ATGCTTCACACGTTTAAACCGAACC 1943  
RESULT 3  
AAC43093  
ID AAC43093 standard; DNA; 2112 BP.  
XX AAC43093;  
XX AC AAC43093;  
XX DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 38002.  
XX DE  
XX DE  
XX KW Hybridization assay; genetic mapping; gene expression control;  
XX KW protein identification; signal transduction pathway; metabolic pathway;  
XX KW promoter; termination sequence; ss.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PF 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 23-MAR-1999; 99US-0123548P.  
XX 25-MAR-1999; 99US-0125788P.  
XX 29-MAR-1999; 99US-0126264P.  
XX 01-APR-1999; 99US-0126785P.  
XX 06-APR-1999; 99US-0127462P.  
XX 08-APR-1999; 99US-0128234P.  
XX 16-APR-1999; 99US-0128714P.  
XX 19-APR-1999; 99US-0129845P.  
XX 21-APR-1999; 99US-0130077P.  
XX 23-APR-1999; 99US-0130449P.  
XX 28-APR-1999; 99US-0130510P.  
XX 30-APR-1999; 99US-0130891P.  
XX 04-MAY-1999; 99US-0131449P.  
XX 05-MAY-1999; 99US-0132407P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132487P.  
XX 11-MAY-1999; 99US-0132487P.  
XX 14-MAY-1999; 99US-0134256P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134768P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136392P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137222P.  
XX 03-JUN-1999; 99US-0137528P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138847P.  
XX 14-JUN-1999; 99US-0139119P.  
XX 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142972P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 18-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144363P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-015753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 42.4%; Score 1038; DB 3; Length 2112;

Best Local Similarity 72.3%; Pred. No. 1.3e-245;

Matches 1363; Conservative 0; Mismatches 520; Indels 3; Gaps 1;

QY 125 CACCAAGTCGAGGTTCCACCGCCAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG 184

Db 193 CACACGGTGGAGGCTCCACAACCTCAACGGTCTTTGAAGTCACTTCAGTACTCAGTGAAG 252

QY 185 GAGACTTTCTTCCCTGATGACCCCTTTGAGGCGAGTTCACAGAACAGCCAGCTTCCCAAGAAG 244

Db 253 GAAACTCTGTTTCAGACGACCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAA 312  
Qy  
Db 245 TTCAATGCTTGGCCCTTCAGTCTTCTCCCAATTTTCGAATGGGCTCCCAATACACCTTT 304  
Db 313 TTTGTGTTAGGCTCAAACTACTCTCTCCGATTTTCGAATGGGACCAACGCTACATCTC 372  
Qy 305 CAGTCTTTGAAAGCTGACCTCATAGCTGGGATACCAATCGCTAGCTTTGGCAATTCCTCAG 364  
Db 373 AAGTCTTTCAAAATCAGATCTCATCGCGGAATCAACCATCGCTAGCTTCGCGCATCCCTCAG 432  
Qy 365 GGCATCAGTTATGTCACCAAGCTCGCAACCTCCCTCCAATTTCTTGACATATATTCGAGCTTT 424  
Db 433 GGCATCAGTTACGCCAAATTCGTAATTCGCCCAATTTCTTGSCCTTTTATTCGAGTTT 492  
Qy 425 ATACCAACATTTGATTTATGGGATGATGGGTAGCTTCGAGGGAATTTGGCAGTGGGACTGTG 484  
Db 493 GTACCGCAATTTGATACGGGTGCTAGGGAGTTCAAGGACTTAGCGGTGGGAACGGTT 552  
Qy 485 GCGTTGGATGCTTCTGATGGGTTTCGATGTTGAGTAATGCGGTTGATCCCAATGAAGAC 544  
Db 553 GCGTTGCGTCTCTGTTGACAGGTGCGATGCTGAGCAAGAAAGTTGATGCTGAGAAAGAT 612  
Qy 545 CCAAGCTTTACCTCCACCTGGCTTTCACAGCTACATTTATTTGCTGGTGTTCAGGCT 604  
Db 613 CCTAAGCTTTACCTTCACCTTTCCTTTCACCGCACTTTTTCGCGCGGCTTCGAGGCC 672  
Qy 605 GCCTTGGGCTGTTTGGGTTGATGCTGGATTTTCTGTCACATGCAACCAATAA 664  
Db 673 TCTCTTGAATTTTCAGTTAGGTTTCATAGTGAATTTCTATCGCATGCAACGATGTA 732  
Qy 665 GGGTTCAATGGAGAGAGCAGCAACGGTGGTGTCTGAGCAACTAAATCGATTCCTGGC 724  
Db 733 GGAATTCATGGAGAGAGCAGCAGCTGGTGTCTGCAACAGCTTAAAGGCTATATTCGGA 792  
Qy 725 CTTGAGCATTTACCCATGAGCTGATATCATATCAGTATGCTGCTGTTTTCACCCAA 784  
Db 793 CTTAAACATTTACAGACTCTACCGATGTTATCTCTGTCATGCGTTCGCTTTCTCCCAA 852  
Qy 785 ACTCATGAGTGGAGTGGGAAAGTGTGTGTAGGATGTCTTTCATTTTCTTCCTCTT 844  
Db 853 ACTCAGAGTGGAGATGGGAAAGTGGGTTCTTGGATGTGGTTTCTATTTCTTCTCTC 912  
Qy 845 AGCAAAAGATCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTGCAAAATCGATGGGCCA 904  
Db 913 TCCACCAAGATTTTCAGCATCAAGAAACCAAAATTCCTTTGGGTGGCGGATGGCTCCT 972  
Qy 905 TTGACGTCCTGTATTTGGGAAGTCTCTTGGTTTATTTTCACTCAACCGGAGAGCAGGT 964  
Db 973 TTGACCTCAGTGATTTTGGGAAGTCTCTTGGTTTACTTCACTCAACGCTGAGAGACATGGT 1032  
Qy 965 GTTGAAGTGATAGGAACTGAAGAAAGGTTTGAATCCACCATCACTCACAAATCTGGTA 1024  
Db 1033 GTTCAAGTGATAGGGACCTGAGAAAGGTTGATCCACTCTCGGTTCTGATCTCATC 1092  
Qy 1025 TTTGTGTGGCTTACATGATACAGCTGTCAAACTGGCAATTTGTCTGGTGGCATCATATCA 1084  
Db 1093 TTTACTTCCCTTACATGTCCACAGCTGTCAAACTGGCCTCATCACTGGCATCATGTCT 1152  
Qy 1085 CTTGCGGAAGGATAGCAGTAGGAGAGCTTTGCAATGTATAAATTTCAATATTTGAT 1144  
Db 1153 CTCGCTGAAGGATAGCAGTGGGAGAGTTTTCGATGTTCGAAGAACTCAACATAGAC 1212  
Qy 1145 GGCACAAAGAGATGATAGCTATTGGGACCATCAACGCTAGTTGGTTCTTTCACTCTTGC 1204  
Db 1213 GGGACAAAGAGATAGCGTTTGGATGATGAAACATCGTTGGTTCTTTCATCATCTTGT 1272  
Qy 1205 TACCTCAACAGGACCAATTTTCGCGTTTCGGCTGTGAACATATAACGCTGGATCAAGACA 1264  
Db 1273 TACCTCAACAGGACCAATTTTCAAGGTTCGCGAGTGAACATCAACGCGGGTTGCAAGACC 1332  
Qy 1265 GCAGCTTCCAAACATTAATGTCACTTTCGAGTAATTTGACATTTATTCCTTGACACC 1324

Db 1333 GCANTGTCCAAATAGTAGTGCGGATTCGGTTATGTTTACACTCTCTTCTTCTCACACCG 1392  
Qy 1325 TTGTTCCATTTACACTCCCTCGTGGTGTCTATCAGCTATTTATCTGTAATCTGCAATGCTTGA 1384  
Db 1393 CTTTTTCACTACACACCACTCGTGGTCTCTCTGCCATCATCATATCCGCAATGCTCGGA 1452  
Qy 1385 CTATAGATTTAAGAGGACCATCCATCTATTTAAGGTTGACAAATTTGACTTTTGGGTG 1444  
Db 1453 CTCAATTTGACTATCAAGCTGCCATCCATCTCTGGAAGTTTGACAAGTTTCGACTTCTCTGTC 1512  
Qy 1445 TGCAATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCT 1504  
Db 1513 TGCAATGAGCGCTTACGTTGGGTGCTATTTGGGCGATGTAGAGATTTGACTCTGCTGTAAGC 1572  
Qy 1505 ATTGTAATATCTGTACTTTCGGGTACTTCTATTTATTGCAAGGCCAAGGACATTCGTTTGTG 1564  
Db 1573 GTGGCGATATCTATAGCGAGTTTGTGCTGTTGTCGAGGCCAANAATCTGCGGTGAAG 1632  
Qy 1565 GGCACCAATTTCCAAATTTCTGTGATATACCGAATTTTGAGCACTATCAAAATGCAAAACAT 1624  
Db 1633 GGAACATATCCAAACAGCATGTATCTATAGGAACACTGAGCAGTACCCTATCATCAAGAACC 1692  
Qy 1625 GTTCTCGGAATGCTAATTTCTAGAGATTTGATGCACCAATTTTACCTTTCGCAATGCGAGCTAT 1684  
Db 1693 GTTCTCGGTATTTCTCATCTTGGAGATTTGATCTCCCATCTACTTTGCTAATGCGAGTTAC 1752  
Qy 1685 TTAAGAGAAAGGATTCACAGGTGGATTTGATGAAGAAAGAAAGAAATTTAAAGCTACAGGG 1744  
Db 1753 TTGCGTGAAGAATCATAGGTGATTTGATGAAGAGGAAGAGAGTTTAAACAATCAGGA 1812  
Qy 1745 GAGACTAGTTTGCAGTATGTTAATTTGATATGAGTGTGTTGGAAACATTTGATACAGT 1804  
Db 1813 GAGAGCAGCTTACAAATATTTACTCGATATGTCAGCTGTTGGTAAATATCGACACAAGC 1872  
Qy 1805 GGAATAAGTATGCTTCAAGAGGTGAAGAAGATTTACAGAGAGAAGAGCTTACAGCTTGT 1864  
Db 1873 GGTATTAGCATGATGTGGAAATTTAAGAAGTCAITTTACAGGAGAGCGTTTAAAGTTGGTA 1932  
Qy 1865 TTGCTCAATCTCTGTAAGTGAAGTGAAGAAACTGAAGAACTGAACAAATCGAAGT---TCCAAAT 1921  
Db 1933 TTGCTCAATCCAAAGAGAGAGGTGCTGAAGAAATTTAACCATGATCCAAATTCATCGGTGAT 1992  
Qy 1922 CATTTAGGAGAAATTTGATCTATCTGACTGTTGAAGAGCGCTTGGAGCATGCAACTTC 1981  
Db 1993 CATTTGGCAAGAGTGGATGTTTAAACGTTAGGAGAAAGCAGTGGAGGCTTGTAGCTAC 2052  
Qy 1982 AATCTACGTGCAAGCAAAACGAAACC 2007  
Db 2053 ATGCTTTCACACGTTTAAACCGAACC 2078

RESULT 4  
ABZ13697  
ID ABZ13697 standard; DNA; 1977 BP.  
XX ABZ13697;  
AC  
XX  
DT 21-JAN-2003 (first entry)  
XX  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1502.  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 24-AUG-2001; 2001WO-US026685.  
XX  
XX 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.



Db 1737 CGGTATTTAGCATGATGTTGGAATTAAGAAAGTCAITTTGACAGGAGCGGTTAAAGTTGGT 1796  
Qy 1864 TTTGGTCAATCTCTGAAGTGAAGTGAATGAAGAACTGAACAAATCGAAGT---TCCAAAA 1920  
Db 1797 ATTGTCAATCCAAAAGGAGAGGTCGTGAAGAAATTAACAGATCCAAATTCATCGGTGA 1856  
Qy 1921 TCATTTAGGAGAAATGGATCTATCTGACTGTTGAAGAGCGCGTTGGAGCATGCAACTT 1980  
Db 1857 TCATTTGGGCAAGAGTGGATGTTCTTTAACGGTAGGAGAGCAGTGGAGGCTTTGTAGCTA 1916  
Qy 1981 CAATCTACGTGCAAGCAAAAGCAACCC 2007  
Db 1917 CATGCTTACACGTTTAAACCGAACC 1943

RESULT 5  
ADA68514  
ID ADA68514 standard; DNA; 1977 BP.  
XX  
AC ADA68514;  
DT 20-NOV-2003 (first entry)  
XX  
DE Arabidopsis thaliana gene, SEQ ID 589.  
XX  
KW Plant; bacterial infection; fungal infection; viral infection; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN W02003000898-A1.  
XX  
PD 03-JAN-2003.  
XX  
PF 22-JUN-2001; 2001WO-IB001105.  
XX  
PR 22-JUN-2001; 2001WO-IB001105.  
XX  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX  
DR WPI; 2003-175290/17.  
XX  
PT Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
XX  
PS Claim 6; SEQ ID NO 589; 899pp; English.  
XX

XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.

SQ Sequence 1977 BP; 513 A; 458 C; 471 G; 535 T; 0 U; 0 Other;  
Query Match 41.68; Score 1018.2; DB 8; Length 1977;  
Best Local Similarity 72.3; Pred. No. 1e-240;  
Matches 1364; Conservative 0; Mismatches 518; Indels 5; Gaps 3;  
Qy 125 CACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAGGTCTCTAAAGTACTCTTTGAAG 184  
Db 58 CACAGGTGGAGGCTCCACAACTCAACCGTCTTGAGTCACTTCAGTACTCAGTGAAG 117

Qy 185 GAGACTTTCTTCCCTGATGACCCCTTTGAGCAGTTCAAGAACAGCCAGCTTCCAGAAAG 244  
Db 118 GAAATCTCTGTTTCCAGACGACCCCTTTTAGACAAATTTAAGAACCAAAATGATCAAGAAAA 177  
Qy 245 TTCATGCTTGGCCTTCAGTCTTCTTCCCATTTTCGAATGGGCTCCCAATACACCTTT 304  
Db 178 TTTGTGTAGGCTCAAAATACCTTCTCCCGATTTTCGAATGGGCAACACGCTACATCTC 237  
Qy 305 CAGTTCTTGAAGCTCACTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAG 364  
Db 238 AAGTTCTTCAAAATCAGATCTCATCGCGGAATCACCATCGTAGTCTCGGCATCCCTCAG 297  
Qy 365 GGCATCAGTTATGCCAAGCTCGCAACCTCCCTCCGAATTTCTTGGACTATATTGAGGTTT 424  
Db 298 GGCATCAGTTATGCCAAGCTTGTAACTTGGCCCCAAATTTCTTGGCCTTTATTTCAGGTTT 357  
Qy 425 ATACCACCATTTGATTTATCGATGAGTGGGTGAGCTCGAGGATTTGGCAGTGGGACCTGTG 484  
Db 358 GTACCGCCATTTGGTATACCGCGTCTAGGAGTTCAAGGACTTAGCGGTGGGAACGGTT 417  
Qy 485 GCGTTTGGATCGCTTCTGATGGGTTCCGATGTTGAGTAATGCCGTTGATCCCAATGAAGAC 544  
Db 418 GCGTTTGGCTCTCTGTTGACAGTGCATGCTGAGCAAGAACTTGATGCTGAGAAAGAT 477  
Qy 545 CCAAAGCTTTACCTCCACCTGGCTTTACAGCTACATTAATTTCGTGGTGT-TTTCAGGC 603  
Db 478 CTTAAGCTTTACCTTCCACCTTGTCTTTCACCGCCACTTTTTTCGCCGCGGTTCTCGAAGCC 537  
Qy 604 TGCCTTGGGTCGTTTATAGTTGGGTTGATCGTGGATTTTCTGTACATCAACCACTAAT 663  
Db 538 TCTCTTGGAAATTTTCAAGGTTAGGTTTCATAGTGGACTTCTATCGATGCAACGATAGT 597  
Qy 664 AGGTTTCATGGGAGGAGCAGCCACCGTGGTGTCTGTCGACAACTAAAAATCGAATTCGTG 723  
Db 598 AGGATTCATGGGAGGAGCAGCGCTGGTGTGAGTCTGCAACAGCTTAAAGGTTATTTTCGG 657  
Qy 724 CTTGAGCATTTTCAACCATGGAGCTGATATCATATCAGTATGCGCTCTGTTTTTCACCCA 783  
Db 658 ACTTAAACATTTTACAGACTCTTACCGATGTTATCTCTGTGATGCGTTCGTTTTCTCCCA 717  
Qy 784 AACTCATAGTGGAGTGGGAAAGTCTGTGTTAGGATGTCTTCTCATTTTCTTCTCTCT 843  
Db 718 AACTCAGAGTGGAGATGGGAAAGTGGCGTCTTGGATGGTTTTCTATCTTCTTCTCT 777  
Qy 844 TAGCACAAGTACTTTCAGCAAAAACGACCAAGTTTTTTTGGGTGTGACAAATGGCGCC 903  
Db 778 CTCACACAGATTTTTCAGCATCAAGAAACCAAAATTTCTTTGGGTGGCGCGATGGCTCC 837  
Qy 904 ATTGACGTCGTTTATTTGGGAAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGACGCG 963  
Db 838 TTTGACCTCAGTGAATTTTGGAAAGTCTCTTGGTTTACTTCTCAGCTCAGAGACATGG 897  
Qy 964 TGTGAAGTATAGGAGAACTGAAGAGGTTTGAATCCACCATCACTCAAAATCTGGT 1023  
Db 898 TGTTCANGTATAGGGGACCTGAAGAAAGGTTGAATCCACTCTCGGTTCTGATCTCAT 957  
Qy 1024 ATTTGTGTCGCTTACATGACTACAGCTGTCAAACTGGCATTTGCTTGGCATCATATC 1083  
Db 958 CTTTACTTCCCTTACATGTCCACAGCTGTCAAACTGGCCTCATCTCAGTGGCATATTGC 1017  
Qy 1084 ACTTGGGAAGGAATAGCAGTAGGAGAGAGCTTTTGAATGTATAAAATTAACAATTTGA 1143  
Db 1018 TCTCGC-GAAGGAGTAGCAGTGGGAGAGGTTTTTGGGATTTTCAAGAACTCAACATAGA 1076  
Qy 1144 TGGCAACAAAGAGATGATAGCTATTGGGACCAAGAGTGTGTTTCTTTCACCTCTTG 1203  
Db 1077 CGGGAACAAAGAGATGATAGCGTTTGAATGATGAACATCGTTTGGTTCTTTCACATCTTG 1136  
Qy 1204 CTACCTTCAACAACAGGACCAATTTTTCGGGTTTCGGGTGTGAACACTATAACGCTGGATGCAAGAC 1263  
Db 1137 TTACCTCAACAACCGGACCAATTTTCAAGGTCGGCAGTGAACACTACACGCGGTTTGAAGAC 1196  
Qy 1264 AGCAGCTTCAACATTTATTAATGTCACTTCAGTAGTAATGTTGACATTTGTTATTTCTCTGACACC 1323



Db 1197 CGCAATGTCACATAGTAGGCGATTGCGGTATGTTACACACTCTCTCTCTCACACC 1256  
Qy 1324 CTTGTTCCATTACACTCCCTGGTGGCTATCAGCTATTATCGTATCTGCAATGCTGG 1383  
Db 1257 GCTTTTCTACTACACACCCTCGTGGCTCTCTGCCATCATCATATCGCAATGCTCGG 1316  
Qy 1384 ACTCATAGATTATGAAGCACCCTCATCTATTATTAAGTTGACAAATTTGACTTTGTGT 1443  
Db 1317 ACTCATGACTATCAAGTGCCTCATCTCTGGAAGTTGACAGTTGCACTTCTCGT 1376  
Qy 1444 GTCCATGAGTGCATACATTGGCTGGTCTTTGGCAGTGTGAAATGGCTTAGTCATAGC 1503  
Db 1377 CTGCATGAGCGCTACGTTGGGTCGTATTCCGCGAGTGTAGAGATTGGACTCGTCTAGC 1436  
Qy 1504 TATTGTAATCTGTACTCTCGGTACTCTTATTATTAAGGCGCCCAAGCATTCGTTTT 1563  
Db 1437 GGTGGCGATATCATACGAGGTTGTGCTGTTGTGTCAGGCCCAAAACTCGCGTGAA 1496  
Qy 1564 GGGCAACATTCCAAATCTCTGATATACCGAAATGTTGACACTATCAAAATGCAAAACA 1623  
Db 1497 GGGAAACATACCAACAGCATGATCTATAGAACACTGACGATCCCATCATCAAGAAC 1556  
Qy 1624 TGTCTCGGAATGCTAATCTAGAGATTGATGACCAATTTACTTTGCCAATGCCAGCTA 1683  
Db 1557 CGTTCTCTGTTATCTCATCTCTGGAGATTGATGCTCCCATCTACTTTGCTAATGCCAGTTA 1616  
Qy 1684 TTTAAGAGAAAGATCACAGGTGGATGATGATGAAGAGAGAAAGATTAAAGCTACAGG 1743  
Db 1617 CTTGCGTGAAGAATCATAGGTGGATGATGAAGAGAGAGAGATTAAACAATCAGG 1676  
Qy 1744 GGAGACTAGTTTTCAGTATCTTATAATGATGATGAGTGTGTTGGAACAATGATACAAG 1803  
Db 1677 AGAGAGCAGTTTCAATATATATCTCGATATGTCAGCTGTTGGTATATTCGACACAG 1736  
Qy 1804 TGAATAAGTATCTGTAAGAGGTGAAGAAGATTACAGAGAGAGAGACTACAGCTTGT 1863  
Db 1737 CGGTATTAGCATGATGTTGGAATTAAGAAAGTCATTGACAGGAGCGTTAAAGTTGTT 1796  
Qy 1864 TTTGGTCAATCTGTAAGTGAAGTATGAGAACTGAGAACTGAACAAATCGAAGT---TCCAAA 1920  
Db 1797 ATTGTCAAATCCAAAAGGAGAGTGTGGAAGAAATTAACAGATCCAAATTCATCGGTGA 1856  
Qy 1921 TCATTTAGGAAGAAATGGAATCTATCTGACTGTTGAAGAGCGGTTGGAGCATGCAACTT 1980  
Db 1857 TCATTTGGGCAAGAGTGGATGTTCTTAACGTTAGGAGCAGTGGAGGCTTGTAGCTA 1916  
Qy 1981 CAATCTACGTGCAAGCAAAACGAACCC 2007  
Db 1917 CATGCTTCACACGTTTAAACCGAACC 1943

RESULT 6

ADAY0002  
ID ADA70002 standard; DNA; 1974 BP.  
XX AC ADA70002;  
XX AC ADA70002;  
DT 20-NOV-2003 (first entry)  
XX DE Rice gene, SEQ ID 3325.  
DE DE Plant; bacterial infection; fungal infection; viral infection; rice;  
XX KW gene; ds.  
KW KW Oryza sativa.  
OS OS WO2003000898-A1.  
XX PN 03-JAN-2003.  
XX PD 22-JUN-2001; 2001WO-IB001105.  
XX PF 22-JUN-2001; 2001WO-IB001105.  
XX XX

PR 22-JUN-2001; 2001WO-IB001105.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.  
DR Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
XX Claim 6; SEQ ID NO 3325; 899pp; English.  
XX The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.  
XX SQ Sequence 1974 BP; 335 A; 602 C; 657 G; 380 T; 0 U; 0 Other;  
Query Match 34.5%; Score 846; DB 8; Length 1974;  
Best Local Similarity 66.4%; Pred. No. 3.3e-198;  
Matches 1231; Conservative 0; Mismatches 620; Indels 3; Gaps 1;  
Qy 131 GTCGAGGTTCCACGCCACAGCGGTTTTCAAGTCTCTAAAGTACTCTTTGAGGAGACT 190  
Db 55 GTGCCGATGCCCGCGCGAAGCGCTTCTCGAGAGCGCTGGGGGGGAACATGAAGGAGACA 114  
Qy 191 TTCTTCCCTGATGACCCCTTTGAGGCGAGTTCAAGAACAGCGCTTCCAAAGATTC--- 247  
Db 115 TTCTTCCCGGACGACCGGTTCAAGGTGGTGGCGGGAGCGGGTGGCGGCGCGCG 174  
Qy 248 ATGCTTGGCGTTTCACTTCTTCCCAATTTTCGAATGGGCTCCCAAAATACACCTTTTACG 307  
Db 175 GCGGGCGGCGCTCCGGTACGTTTCCGTTTCATGAGAGTGGGCGCGCTGCTACACCTCGGC 234  
Qy 308 TTCTTGAAGCTGACCTCATAGCTGCATCACCATCGTAGCTTGCCCATTTCTCAGGGC 367  
Db 235 ACCCTCAAGTCCGACCTCATCGCGGCATCACATTCGAGCGCTCGCCATCCCCAGGGC 294  
Qy 368 ATCAGTTATGCCAAGCTCGCCAACTCCCTCCCAATTTCTTGGACTATATTCGAGCTTTATA 427  
Db 295 ATCAGTACGCCAAGCTCGCCAACTCCCTCCCGCTCTCTATTCGAGCTTCGTG 354  
Qy 428 CCACCATTTATTTAGCGATGATGGTAGCTCGAGGATTTGGCAGTGGGAGCTGTGGCG 487  
Db 355 CCGCGCGTGGTACGCGATGATGGGAGCTCGAGGAGCTCGAGGAGCTCGGCGTGGGACG 414  
Qy 488 GTTGGATCGCTTCTGATGGGTTGATGTTGAGTAATGCCGTTGATCCCAATCAAGACCCA 547  
Db 415 GTGGCGTCTGCTGATCGGGTGCATGCTGACGAGGAGGTGTCGGCGCGGAGGACCCG 474  
Qy 548 AAGCTTTTACCTCCACTCGCTTTTACAGCTACATTTTGTGCTGGTGTGTTTTTCAGGCTGCC 607  
Db 475 GCGCTGTACTGCAGCTCGCGCTCACCGCCACTTCTTCGCCGCGGTGTTTCAGGCGCTG 534  
Qy 608 TTGGGTCTGTTAGGTTGGGTTGATCGTGGATTTTCTGTACATGCACACCAATATAGGG 667  
Db 535 CTGGGGGTCTTCAGGCTGGGTTTCATCGTGGACTTCTCTGTCGACGCCACCATCGTCGG 594  
Qy 668 TTCATGGGAGGAGAGCCACGCTGTGTCTGTCGACGAACTAAATTCGATTTCTTGGCTT 727  
Db 595 TTCATGGGCGCGCGCCACCGGTGTGTGCTTGCATGACGAGCTCAAGGGCATGTTGCGCCTC 654



Qy 728 GAGCATTTACCCATCGAGCTGATATCATATCAGTATGCGCTCTGTGTTTTCACCCAACT 787  
Db 655 GACCACCTTACACCGCCACCGACCTCGTCTCCGTCATGAGCTCCGCTTTCTCCCAACC 714  
Qy 788 CATGAGTGGAGCTGGGAAAGTGTGTAGGATGTGTCTTCACTTTCTCTCTCTTAGC 847  
Db 715 CACCTCTGGCGATGGGAGAGCGTGTCTATGGCTGCGGCTTCTTCTTCTCTCTCATC 774  
Qy 848 ACAAGATACCTTACGAAAAAACAAGGTTTTTTTGGGTGTGAGCAATGGGCCCATG 907  
Db 775 ACCCGCTTCTTACGAAAGAGGAGGCAAGGTTCTTCTGGGTATCTGCAGCTGCGCCATTG 834  
Qy 908 ACGTCCGTATATTGGGAAGTCTCTGTGTTTATTCTACCTCAGCCGCGAGAGCAGGTGTT 967  
Db 835 GCGTCTGTATCATCTCGGAGGCTCTCGTGTACTCACACATGTCTGAAACCATGGCAAT 894  
Qy 968 GAAGTGTATGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAAATCTGTATTT 1027  
Db 895 CAAGTGATTGGTTACCTGAAGAAGGCTGAAACCCACCGTCTGCCACAGGCTGAACCTTC 954  
Qy 1028 GTGTCCCTTACATGACTACAGCTGTCAAACTGGCAATGTGCTGGATCATATCACTT 1087  
Db 955 TCGTCCGCGTACATGATGCTGCGCTGAAGACCGGATCATCACTGGCGTCAATGGCCCTC 1014  
Qy 1088 GCGGAAGGATAGCAGTAGGAGAGCTTTGCAATGTATAAAATTACAATATTGATGGC 1147  
Db 1015 GCTGAGGGATTGCGGTAGGAAGAGCTTTGCAATGTTCAGAAATTACCAATCGACGGA 1074  
Qy 1148 AACAAAGAGATGATAGCTATTGGGACCATGAAGTGTGTTCTTTCACTCTTGCTAC 1207  
Db 1075 AACAAAGAGATGATGCGGTTTGGAAAGATGAACATTTGTTGGATCACTCACTCTTGCTAC 1134  
Qy 1208 CTCAACAAGACCATTTTTCGCTGCTGCTGTAAGTATTAACGCTGGATGCAAGACAGCA 1267  
Db 1135 CTCAACACGCTCGCTTCTCGAGCTCGCGCTCAACTACAAAGCGCGCTGCAAGACGCG 1194  
Qy 1268 GCTTCCAACTATTAATGTCACTTGCAGATATGTTGATGTTTATTCCTGACACCTTG 1327  
Db 1195 ATGTCGAACGTGATCATGTGCGTGGCGTGTATGATCATCGCTGCTGTGTTCTTGACCGCGCTG 1254  
Qy 1328 TTCCATTACACTCCCTCGTGTGCTATCATGATTTATTCGATGTCGCAATGCTTGGACTC 1387  
Db 1255 TTCCACTACAGCGCTGTGTGCTGTGCGGATCATCATGTGCGGATGCTGGGCTG 1314  
Qy 1388 ATAGATTATGAAGCAGCATCACTATTTAAGTTTGACAAATTTGACTTTGTGTGTC 1447  
Db 1315 ATCGACTACCGCGCGCTCCACTGTGGCAGGTGGACAAAGTGGACTTCTGCTGCTGC 1374  
Qy 1448 ATGAGTCATACATGCGGTGCTTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCTATT 1507  
Db 1375 CTGCGCGCTTACCTCGCGGTGCTTCTCGGCAAGCTCAGATCGGCTCGTGGCGCTC 1434  
Qy 1508 GTAATATCTGTACTTCGGTACTTCTATTATTCGAAAGGCAAGGACATTCGTTTGGGC 1567  
Db 1435 GGGATCTCCATCTCCGCTGCTGCTGTGTCGCGCGCGGAGGACGAGCTGCTCGG 1494  
Qy 1568 AACATTCCAAATCTGTATATACCGAAATGTTGAGCACTATCAAAATGCAAAACATGTT 1627  
Db 1495 AACATCCCAACAGATGATCTACCGCGGATGACCAAGTACACCGCGCGAGAGGCTC 1554  
Qy 1628 CTTGAATGCTAATTTCTAGATTTGATGCAACCAATTTACTTTGCAATGCGAGTATTTA 1687  
Db 1555 CCGCGCGTCTGCTGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1614  
Qy 1688 AGAGAAAGGATCACAGGTGATGATGAAGAAGAAAGAAATTAAGCTTACAGGGGAG 1747  
Db 1615 CGTGAGAGATTCGCGCGTGTGATGACGACGAGGAGCCAGTGCAGAGAGAGAGGGGAG 1674  
Qy 1748 ACTAGTTTGCAGTATGTTTATTAATTTGATAGTGTCTGTTGGAAACATTTGATACAGTGA 1807  
Db 1675 ATGGGCATTCAGTACGTCGCTCTCGACATGGTGGCTGCGAGCATCGACACGCGGG 1734  
Qy 1808 ATAAGTATGTTGAAGGTGAAGAAGATTACAGAGAGAGAGACTACAGCTGTTGTTTG 1867

Db 1735 ACAGCATGCTGATGAATCAAGAGACCTCGACAGGAGGCGCTTCAGATTGTGTTG 1794  
Qy 1868 GTCAATCTCTTAAGTGAAGTGAAGAAACTGAAACAAATCGAAGTTCCAAATCAATTTA 1927  
Db 1795 GCGAACCCGGGAGCGAGATCATGAAGAAGTTGGACAGCTCCAAGGTGCTTGAGCGCATC 1854  
Qy 1928 GGGAAAGAAATGATCTATCTGACTGTTTGAAGAGCGCTTGGAGCATGCAACTTC 1981  
Db 1855 GGCCATGAGTGGATCTTCCCACGCTGGGCGAGCGCTGCGGAGTGCAGCTTC 1908

RESULT 7  
AAZ50482  
ID AAZ50482 standard; cDNA; 1981 BP.  
XX  
AC AAZ50482;  
XX  
DT 23-MAY-2000 (first entry)  
XX  
XX Corn sulphate permease-2 cDNA clone.  
DE Sulphate Permease; sulphate assimilation protein; corn; contig; probe;  
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;  
KW antibody; screen; ss.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 2..1741  
FT /tag= a  
FT /product= "Corn sulphate permease-2"  
FT /note= "Derived from contig composed  
FT crln.pk0015.a2, p0006.cbyvs25rb, p0072.comhc25r,  
FT p0091.cward29r and p0092.chwat43r"  
XX  
PN WO200004154-A2.  
XX  
XX 27-JAN-2000.  
XX  
XX 13-JUL-1999; 99WO-US015810.  
XX  
XX 14-JUL-1998; 98US-0092833P.  
PR (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
XX Allen SM, Falco SC, Thorpe CJ;  
XX  
XX WPI; 2000-195025/17.  
XX P-PSDB; AAY44936.  
XX  
XX Nucleic acid fragments encoding sulfate assimilation proteins in plants  
XX PT and seeds useful as probes for isolating cDNAs and genes encoding  
XX PT homologous proteins, in producing transgenic plants.  
XX  
XX Claim 3; Page 41; 79pp; English.  
XX  
XX The present sequence is a cDNA clone encoding corn sulphate permease, a  
XX sulphate assimilation protein. This sequence is from a contig composed of  
XX clones crln.pk0015.a2, p0006.cbyvs25rb, p0072.comhc25r, p0091.cward29r,  
XX and p0092.chwat43r from crln, p0006, p0072, p0091, and p0092 libraries,  
XX respectively. This sequence is used as a probe to isolate other plant  
XX sulphate assimilation proteins, for genetic and physical mapping of  
XX related genes and as markers of traits linked to the gene. This is useful  
XX for plant breeding and to construct chimeric genes, used to create  
XX transgenic plants with altered levels of sulphate permease. The sulphate  
XX permease peptides are useful for producing antibodies, that are used to  
XX screen and isolate cDNA clones  
XX  
SQ Sequence 1981 BP; 388 A; 621 C; 612 G; 360 T; 0 U; 0 Other;  
Query Match 31.2%; Score 765.2; DB 3; Length 1981;  
Best Local Similarity 65.8%; Pred. No. 2.9e-178;

Matches 1112; Conservative 0; Mismatches 578; Indels 0; Gaps 0;			
Qy	313	GAAGAGTACCTCATAGCTGGCATCACCATCGTAGCTTGGCCATTCCTCAGGCGATCAG	372
Db	7	GGAGTCGACCTGATCGCCGGCATCACCATCGCAGCTCGCCATCCCGAGGCGATCAG	66
Qy	373	TTATGCCAAGCTGCCAACCTCCCTCCAAATCTTTGGACTATATTCGAGCTTTATACACC	432
Db	67	CTAGCCCAAGCTGCCAACCTCGCCCGCGTCTCGGACTCTACTCGAGCTTCGTGCGGC	126
Qy	433	ATTGATTTATGCGATGATGGTAGCTCGAGGGATTTGGCAGTGGGAGCTGTGCGGTTGG	492
Db	127	GCTGGTGTACCGCTGATGGGGAGCTCCAAAGACCTGGCGGTGGGACGGTGGCGTGGC	186
Qy	493	ATCGCTTCTCATGGGTTGAGTGTGAGTAATGCGTTGATCCCAATGMAAGCCCAAGCT	552
Db	187	GTGCTGCTCATCAGCTTCCATGCTCGGCACGAGGTGTGCCACGGAGAACCCCGTGT	246
Qy	553	TTACCTCCACCTGGCTTTACAGCTACATTTATTTGCTGTGTTTTCAGGCTGCCTTGGG	612
Db	247	CTACCTGCACCTGCCTTTCACGCCACCTTCTTCGCGGGTCTTCCAGGCTCGCTCGG	306
Qy	613	TCGTTTAGTGGGGTTGATCGTGAATTTCTGTCAATGCAATGCAACCATATAGGGTTCA	672
Db	307	CCTCCTCAGGTTGGGTTTCACTCGTGACCTGCTGTGCGACGCGACGATCGTGGGGTTCA	366
Qy	673	GGGAGGAGCAGCAGCTGGTGTCTGCAGCAACTAAATCGATTTCTTGGCCCTTGAGCA	732
Db	367	GGCCGGCGGCGACGGTGTGTGCTGCAGCAGCTGAAGGGCATGCTGGGCCCTCGTCCA	426
Qy	733	TTTCACCCATGGAGCTGATATCATATCAGTGATGCGCTCTGTGTTTCAACCCAACTCATGA	792
Db	427	CTTCAACCACTCCACCGACGTCGCTCCGTCATGGAATCGTCTTCAGCCAGACACCA	486
Qy	793	GTGAGGTGGGAAGTCTGTGTTAGGATGTGTCTTCAATTTTCTTCTCTTTAGCAAG	852
Db	487	GTGGCGTGGGAGAGCTCTGCTGCGCTGCGGCTTCTCTTCTCTCTCTGCTCACCG	546
Qy	853	ATACTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTACAGCAATGCGCGCATTTGACGTC	912
Db	547	CTTCAATCAGCAAGGGTCCAAAGCTTCTGGAATCTCCGCGCGCGCGCTTTGACGTC	606
Qy	913	CGTTATATTGGGAAGTCTCTTGGTTTATTTCACTCACGCGGAGAGCACGTGTGGAAGT	972
Db	607	CGTCGTCTCGGGAGCGTCTGCTGTGTACCTCACGACGCTGAAACCAACGCGATCGAAGT	666
Qy	973	GATAGGAACTGCAAGAGGGTTGAATCCACCATCACTCACAATACTGTGTAATTTGTGTC	1032
Db	667	GATCGGTTACTGAAGAAAGGCGCTGAATCCACCGTCGGTGACAAAGCTTCAATTTCTCACC	726
Qy	1033	GCCTTACATGACTACAGCTGTCAAAATGCGCATTTGTGTTGGCATCATATCACTTCGCGGA	1092
Db	727	GCCTTACATGATGCTCGCGCTCAAGACTGGGATCATCACCGGCGTCAITGCCCCTCGCCGA	786
Qy	1093	AGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAATTAACAATTTGATGGCAACAA	1152
Db	787	AGGAATCGCGTGGGGAGAGCTTCGCCATGTTCAAGACTACCAATGACGAGCAACAA	846
Qy	1153	AGAGATGATAGCTATTGTTGGGACCATGAACGTTAGTGTGTTCTTCACTCTTGTCTACTCAC	1212
Db	847	GGAGATGATCGCATCGGACGATGAACGCTCTGGGCTGCTCAGCTGCTGCTACCTGAC	906
Qy	1213	ACAGAGACCAATTTTCGGTTTCGGCTGTGAACTATAAGCTGGATGAAGACAGCAGCTTC	1272
Db	907	CACGGGGCCCTTCTCGCGCTCCCGCGTGAACACTACAACCGCGGTGAGGACGGCCATGTC	966
Qy	1273	CAACATTATAATGTCACTTGCAGTAATGTGACATTTGTTATTCCTGACACCCCTTGTTC	1332
Db	967	GAACGTGGTCTATGCTGCTGCGGTGATGGTCAAGCTGCTGTTCTTCAACGCGCTGTTCA	1026
Qy	1333	TTACACTCCCTCGTGGTGTATCAGCTATTATCGTATCTGCAATGCTTGGACTCATAGA	1392
Db	1027	CTACACCGCGCTGGTGTGCTGTGCGGATCATCGTCTCGCGATGCTGGGCTTGGTCA	1086

Qy	1393	TTATGAAGCAGCATCCATCTATTTAAGGTTGACAAATTTGACTTTGTGGTGTGCATGAG	1452
Db	1087	CTTGGGGCGCGGTGCACCTGTGGCGGTGCACAAAGTTCGACTTCTGCTGTGCGCCGG	1146
Qy	1453	TGCATACATTGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTTAGTCATAGCTATTGTAAT	1512
Db	1147	CGGTACCTGGGGTGTCTTCGGCAGCTTCGAGGTGCGSCTGGTCTGTCGCGTGCCTGT	1206
Qy	1513	ATCTGTACTTCCGGTACTTCTTATTTATTTCGAAGGCCAAGACATTCGTTTTGGGCAACAT	1572
Db	1207	CTCCCTGCTCCCGTCTCTGCTGTCGCCCGCGCCAGGACACGGTCTCTCGGCAACAT	1266
Qy	1573	TCCAAATTCGTGATATACCGAAATGCTTGAGCACATATCAAAATGCAAAACATGTTCCCTGG	1632
Db	1267	CCCGGCACCATGTTGTACCGAGGATGGACAGTACCGCGCGCGCACACGGTGCCTGG	1326
Qy	1633	AAATGCTAAATTTAGAGATTGATGCACCAATTTACTTTGCAATGCGCAGCTATTTAAGAGA	1692
Db	1327	CGTCTCGTGTGCGGTGCGACGCGCCCTCTACTTTCGCCAACGCCAGCTACTCTGGAGA	1386
Qy	1693	AAGGATCAACAGTGTGATTAAGAGAAAGAAATTAAGCTACAGGGGAGACTAG	1752
Db	1387	GAGGATCTCGCGTGTGATCGACGACGAGGAGGCGCACCAAGAGCCAGGGCGAGATGGG	1446
Qy	1753	TTTGCAATGTTATAATTTGATATGAGTGTCTGTTGAAAACATTGATACAAGTGGATTAAG	1812
Db	1447	CGTGGGTACGTTGTCTCGACATGGTGTCCATCGGTAGCATCGACAGCGGACGAG	1506
Qy	1813	TATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGAGCTACAGCTTTGTTTGTCAA	1872
Db	1507	CATGCTGGACGAGCTCAACAAGTCTTGGACAGGAGGGAATGCAGATCGTGTGCGGAA	1566
Qy	1873	TCTGTAGTGAAGTGAAGAAACTGAAACAAATCGAAGTTCGAAATCATTTAGGGA	1932
Db	1567	CCCGGACGCGAGATCATGAAGAAGTGGACAGCTCCAAGGTGCTGGAGCAGATCGGCCA	1626
Qy	1933	GAAATGGATCTATCTGACTGTTGAAGAGCGCTTGGAGCATGCAACTTCAATCTACGTGC	1992
Db	1627	CGAGTGGGTGTTCCCGACGCTGGCGGAGCGGTGGCTGTGGACTACGTCTGCTGCACTC	1686
Qy	1993	AAGCAAAACG 2002	
Db	1687	GCACAAGCCG 1696	

RESULT 8

ID	AAZ50488	standard; cdNA; 2067 BP.
XX	AAZ50488;	
AC	AAZ50488;	
XX	23-MAY-2000	(first entry)
DT	23-MAY-2000	(first entry)
XX	Soybean sulphate permease-1 cdNA clone.	
DE	Sulphate Permease; sulphate assimilation protein; soybean; probe;	
KW	mapping; marker; plant breeding; chimeric gene; transgenic plant;	
KW	antibody; screen; ss.	
XX	Glycine max.	
XX	Key	Location/Qualifiers
FT	CDS	2..1867
FT	/product= a	
FT	/note= "Soybean sulphate permease-1"	
XX		/note= "Derived from clone ses2w.pk0031.b3"
PN	WO200004154-A2.	
XX	27-JAN-2000.	
XX	13-JUL-1999;	99WO-US015810.
PF		



Db	1688	GAAGCTTAGAAGATGCTGGAGAGAGATCACTTGAAGTTTGTGCTGGCAAAATCCTGTTGGA	1747	PR	08-JUN-1999;	99US-0138094P.
				PR	10-JUN-1999;	99US-0138540P.
Qy	1883	GAAGTGATGAAGAACTGACAAATCGAAGTTCAAAATCATTTAGGGAAGAAATGGATC	1942	PR	10-JUN-1999;	99US-0138847P.
				PR	14-JUN-1999;	99US-0139119P.
Db	1748	AATGTGAATGAAGAAATTCATAGTCAAAATTTGGATTCCTTTGGATTTAAAGGAGTC	1807	PR	16-JUN-1999;	99US-0139452P.
				PR	16-JUN-1999;	99US-0139453P.
Qy	1943	TATCTGACTCTTGAAGAGGCCGT	1965	PR	17-JUN-1999;	99US-0139492P.
				PR	18-JUN-1999;	99US-0139454P.
Db	1808	TATCTCACAGTGGGAGAACTGT	1830	PR	18-JUN-1999;	99US-0139455P.
				PR	18-JUN-1999;	99US-0139456P.
				PR	18-JUN-1999;	99US-0139457P.
RESULT 9				PR	18-JUN-1999;	99US-0139458P.
AAC42229				PR	18-JUN-1999;	99US-0139459P.
ID AAC42229 standard; DNA; 2006 BP.				PR	18-JUN-1999;	99US-0139460P.
AC AAC42229;				PR	18-JUN-1999;	99US-0139461P.
AC				PR	18-JUN-1999;	99US-0139462P.
XX				PR	18-JUN-1999;	99US-0139463P.
DT 17-OCT-2000 (first entry)				PR	18-JUN-1999;	99US-0139750P.
XX				PR	18-JUN-1999;	99US-0139763P.
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 34765.				PR	21-JUN-1999;	99US-0139817P.
XX				PR	22-JUN-1999;	99US-0139899P.
KW Hybridisation assay; genetic mapping; gene expression control;				PR	23-JUN-1999;	99US-0140353P.
KW protein identification; signal transduction pathway; metabolic pathway;				PR	23-JUN-1999;	99US-0140354P.
KW promoter; termination sequence; ss.				PR	24-JUN-1999;	99US-0140695P.
XX				PR	28-JUN-1999;	99US-0140823P.
OS Arabidopsis thaliana.				PR	29-JUN-1999;	99US-0140991P.
XX				PR	30-JUN-1999;	99US-0141287P.
XX				PR	01-JUL-1999;	99US-0141842P.
XX				PR	01-JUL-1999;	99US-0142154P.
PN EP1033405-A2.				PR	02-JUL-1999;	99US-0142055P.
PD				PR	06-JUL-1999;	99US-0142390P.
XX				PR	08-JUL-1999;	99US-0142803P.
PF				PR	09-JUL-1999;	99US-0142920P.
XX				PR	12-JUL-1999;	99US-0142977P.
25-FEB-2000; 2000EP-00301439.				PR	13-JUL-1999;	99US-0143542P.
XX				PR	14-JUL-1999;	99US-0143624P.
25-FEB-1999; 99US-0121825P.				PR	15-JUL-1999;	99US-0144005P.
PR				PR	16-JUL-1999;	99US-0144085P.
05-MAR-1999; 99US-0123180P.				PR	16-JUL-1999;	99US-0144086P.
PR				PR	19-JUL-1999;	99US-0144325P.
09-MAR-1999; 99US-0125788P.				PR	19-JUL-1999;	99US-0144331P.
PR				PR	19-JUL-1999;	99US-0144332P.
25-MAR-1999; 99US-0126264P.				PR	19-JUL-1999;	99US-0144333P.
PR				PR	19-JUL-1999;	99US-0144334P.
29-MAR-1999; 99US-0126785P.				PR	19-JUL-1999;	99US-0144335P.
PR				PR	20-JUL-1999;	99US-0144352P.
01-APR-1999; 99US-0127462P.				PR	20-JUL-1999;	99US-0144632P.
PR				PR	20-JUL-1999;	99US-0144884P.
06-APR-1999; 99US-0128234P.				PR	21-JUL-1999;	99US-0144814P.
PR				PR	21-JUL-1999;	99US-0145086P.
08-APR-1999; 99US-0128714P.				PR	21-JUL-1999;	99US-0145088P.
PR				PR	22-JUL-1999;	99US-0145087P.
16-APR-1999; 99US-0129845P.				PR	22-JUL-1999;	99US-0145089P.
PR				PR	22-JUL-1999;	99US-0145192P.
19-APR-1999; 99US-0130077P.				PR	23-JUL-1999;	99US-0145145P.
PR				PR	23-JUL-1999;	99US-0145218P.
21-APR-1999; 99US-0130449P.				PR	23-JUL-1999;	99US-0145224P.
PR				PR	26-JUL-1999;	99US-0145276P.
23-APR-1999; 99US-0130510P.				PR	27-JUL-1999;	99US-0145913P.
PR				PR	27-JUL-1999;	99US-0145918P.
23-APR-1999; 99US-0130891P.				PR	27-JUL-1999;	99US-0145919P.
PR				PR	28-JUL-1999;	99US-0145951P.
28-APR-1999; 99US-0131449P.				PR	02-AUG-1999;	99US-0146386P.
PR				PR	02-AUG-1999;	99US-0146388P.
30-APR-1999; 99US-0132048P.				PR	02-AUG-1999;	99US-0146389P.
PR				PR	03-AUG-1999;	99US-0147038P.
04-MAY-1999; 99US-0132407P.				PR	04-AUG-1999;	99US-0147204P.
PR				PR	04-AUG-1999;	99US-0147302P.
05-MAY-1999; 99US-0132484P.				PR	05-AUG-1999;	99US-0147192P.
PR				PR	05-AUG-1999;	99US-0147260P.
05-MAY-1999; 99US-0132485P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
06-MAY-1999; 99US-0132486P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
06-MAY-1999; 99US-0132487P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
07-MAY-1999; 99US-0132863P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
11-MAY-1999; 99US-0134256P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
14-MAY-1999; 99US-0134218P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
14-MAY-1999; 99US-0134219P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
14-MAY-1999; 99US-0134221P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
14-MAY-1999; 99US-0134370P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
18-MAY-1999; 99US-0134768P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
19-MAY-1999; 99US-0134941P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
20-MAY-1999; 99US-0135124P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
21-MAY-1999; 99US-0135353P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
24-MAY-1999; 99US-0135629P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
25-MAY-1999; 99US-0136021P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
27-MAY-1999; 99US-0136392P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
28-MAY-1999; 99US-0136782P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
01-JUN-1999; 99US-0137222P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
03-JUN-1999; 99US-0137528P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
04-JUN-1999; 99US-0137502P.				PR	06-AUG-1999;	99US-0147303P.
PR				PR	06-AUG-1999;	99US-0147303P.
07-JUN-1999; 99US-0137724P.				PR	06-AUG-1999;	99US-0147303P.

PR	09-AUG-1999;	99US-0147493P.	QY	107	AACAACCTTTGAGAGAGTGACCAAGTCGAGTTTCCACCGCCACAGCGGTTTTTCAAGTCT	166
PR	09-AUG-1999;	99US-0147935P.	DB	102		
PR	10-AUG-1999;	99US-0148171P.	DB	102	AACCAATTAATGAGGTTGCACAGGTGGTTGCTCCGCCGATAGAGACGGTGGCAAG	161
PR	11-AUG-1999;	99US-0148319P.	QY	167	CTAAAGTACTCTTTGAAGGAGACTTTCTCCCTGATGACCTTTTGAGGAGTTCAGAAC	226
PR	13-AUG-1999;	99US-0148341P.	DB	162		
PR	13-AUG-1999;	99US-0148565P.	DB	162	CTAAAGACAAACTGAAGGAGACTTTCTCCCGACGATCCCTTAAGACAGTTTCAGAGGA	221
PR	16-AUG-1999;	99US-0149388P.	QY	227	AAGCCAGCTTCCAAGAAAGTTCAATGCTTGGCTTTCAGTTCCTTCTTCCCCTTTTTCGAATGG	286
PR	18-AUG-1999;	99US-0149175P.	DB	222	CAACCAAAACCGTACCAAGCTCATACGCGCGCTCAATACATTTTTCCTTCCATCCCTCAATGG	281
PR	20-AUG-1999;	99US-0149722P.	QY	287	GCTCCCAATACACCTTTTCAAGCTGACCTCATAGCTGGCATCACCATCGCT	346
PR	20-AUG-1999;	99US-0149723P.	DB	282	TGTCCTGAGTACAGCTTTCAGCTTCAAAATCCGACGTCGTTCAGGTCTTTACATCGCT	341
PR	23-AUG-1999;	99US-0149902P.	QY	347	AGCTTGGCCATTCCTCAGGGCATCAGTTATGCAAGCTGCCCAACCTCCCTCCCAATTCCT	406
PR	25-AUG-1999;	99US-0150566P.	DB	342	AGTTTAGCTATTCTCAGGGGATAGTTACGCGAGCTAGCAAAATTTGCCACCAATCGTT	401
PR	26-AUG-1999;	99US-0150884P.	QY	407	GGACTATATTCGAGCTTTTATACCAATTTGATTCGATGATGGTACGTCGAGGGAT	466
PR	27-AUG-1999;	99US-0151065P.	DB	402	GGTCTATACCTCGAGCTTCGTGCCACCGTTGGTTTATGCGGTCTTTGGGAGCTCAAGAGAT	461
PR	27-AUG-1999;	99US-0151066P.	QY	467	TTGGGAGTGGGACTGTGGGGTTGGATCGCTTCTGATGGGTTGATGTTGAGTAATGCC	526
PR	27-AUG-1999;	99US-0151080P.	DB	462	CTAGCAGTGGGACCACTCTCCATAGCTCGTTAACTTTAGGATCCATGCTAAGGCAACAA	521
PR	30-AUG-1999;	99US-0151303P.	QY	527	GTTGATCCCAATGAAGACCCAAAGCTTTTACCTCCACCTGGCTTTTCACAGCTACATATTT	586
PR	31-AUG-1999;	99US-0151438P.	DB	522	GTATCTCCCGTAGACGATCTGTCTCTTTCTACAGCTAGCTTCTCTCTTCTACCTCTTT	581
PR	07-SEP-1999;	99US-0151930P.	QY	587	GCTGTGTTTTTTCAGGCTGCTTGGGTCTGTTTAGTGTGGGTGATCGTGGATTTTCTG	646
PR	07-SEP-1999;	99US-0152363P.	DB	582	GCTGTCTCTTTCAAGCCCTCTCTTGAATCTCTCAGGCTGGGATTTATAATAGACTTTCTA	641
PR	10-SEP-1999;	99US-0153070P.	QY	647	TCATATGCAACCAATATAGGTTTCATGGGAGGAGGAGCCACGGTGGTGTGCTGAGAGCA	706
PR	13-SEP-1999;	99US-01533758P.	DB	642	TCAAAAGCGACGCTAATAGGGTTTATGGGTGGAGAGCCATAATTTGATACACTCCAAACAG	701
PR	13-SEP-1999;	99US-01534018P.	QY	707	CTAAATCGATTCTTGGCTTGAGCATTTTCAACCATGAGCTGATATCATATCAGTGATG	766
PR	16-SEP-1999;	99US-0154039P.	DB	702	CTAAGGGTCTGTTGGGATTACTCAATTTCAAAAGCATATGAGTGTAGTCCCTGTCTC	761
PR	20-SEP-1999;	99US-0154779P.	QY	767	CGCTCTGTTTTTCAACCAAACTCATGAGTGGAGTGGGAAAGTCTGTGTTAGGATGTGC	826
PR	04-OCT-1999;	99US-0155486P.	DB	762	TCCTCTGTTTTTCAACACACACAGAGTGTCTATGSCAAACAAATTTGTAAGGAGTTGC	821
PR	06-OCT-1999;	99US-0155753P.	QY	827	TTCAATTTCTCTCTCTTAGCAGAGATACCTTCAGCAAAAACGACCAAGGTTTTTTTGG	886
PR	07-OCT-1999;	99US-0156458P.	DB	822	TTCTTGCTCTTCTTGTCTCGACACGCTCACCTCAGCATGAAGAAGCGAAGCTGTTTTGG	881
PR	08-OCT-1999;	99US-0156596P.	QY	887	GTGTACAGCAATGGCGCATTTGACGCTCGTTATATTCGGGAGTCTCTTGGTTTATTTCACT	946
PR	12-OCT-1999;	99US-0157865P.	DB	882	GTCTACGCGGAGCACCTTCTTCCGTTATGCTCTACATTTCTGTCTTGTCTTGTGTTTC	941
PR	13-OCT-1999;	99US-0158232P.	QY	947	CAGCGCGAGAAGCACGGTGTGGAAGTGATAGGAGAACTGAAGAAGGGTTTGAATCCACCA	1006
PR	14-OCT-1999;	99US-0158369P.	DB	942	AGAGCGGAGCGCTCAGGAATCAGCGTCTCGGGAATACCAGAAAGTTTGAATCCACCG	1001
PR	18-OCT-1999;	99US-0159293P.	QY	1007	TCACTCAGAAATCTGGTATTTGTGTCGCTTACATGACTACAGCTGTCAAAACTGGCAT	1066
PR	21-OCT-1999;	99US-0160741P.	DB	1002	TCTTGGAAATGCTTTCAGTTTCAAGTGTCTATCTCGACTCTCGTCCCAAAACCGACTC	1061
PR	21-OCT-1999;	99US-0160767P.	QY	1067	GTGCTGGCATCATATCCTTTCGGGAAGGAATAGCAGTAGGAGAGCTTTGCAATGTAT	1126
PR	21-OCT-1999;	99US-0160768P.	DB	1062	GTACCGGAATCGTCTCCCTCAGGAAGGAATCGCAGTGGGAGAAACAATTTGCGCGCTA	1121
PR	21-OCT-1999;	99US-0160815P.	QY	1127	AAAAATTACAATTTGATGGCAACAAAGATGATAGCTATTTGGGACCACTGAACGTAGTT	1186
PR	22-OCT-1999;	99US-0160981P.	DB	1122	AGAACTACCGTAGATGGAACAAAGAGATGATGCCATTGGTCTGATGAACGTAGTA	1181
PR	25-OCT-1999;	99US-0161404P.	QY	1187	GGTCTCTTCACTCTTGTGCTACCTCAACAGGACCAATTTTCGCGTTCCGCTGTGAACTAT	1246
PR	25-OCT-1999;	99US-0161405P.				
PR	25-OCT-1999;	99US-0161406P.				
PR	26-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.				
PR	26-OCT-1999;	99US-0161361P.				
PR	28-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161922P.				
PR	28-OCT-1999;	99US-0161993P.				
PR	29-OCT-1999;	99US-0162142P.				

Query Match 27.0%; Score 662.4; DB 3; Length 2006;  
 Best Local Similarity 60.1%; Pred. No. 6.8e-153;  
 Matches 1120; Conservative 0; Mismatches 741; Indels 3; Gaps 1;



||||| 547 ATAGGGTTATGGTGAGCAGCCATAATTTGTAATCACTCAACAGCTAAAGGGTCTGCTT 606  
||||| 722 GGCTTTGAGCATTTTCAACCATGGAGCTGATATCATATCACTGATGCGCTCTGTTTTCAAC 781  
||||| 607 GGGATRAACTCATTTCAACAGCATATGAGTGTAGTCCCTCTCTCTCTCTGTTTTCCAA 666  
||||| 782 CAAACTCATGATGAGGAGTGGAAAGTCTGTTGATGAGTGTCTTTCATTTCTCTCTC 841  
||||| 667 CACACCAACGAGTGGTCAATGCGCAACAATTTGATGGAAGTGTGCTCTCTCTCTCTG 726  
||||| 842 CTTAGCACAAGATACCTTCAACAAAACGACCAAGGTTTTTTGGGTGTCAGCAATGGCG 901  
||||| 727 CTCTGACACGTCACTCAGCATGAAGACCGAAGCTGTTTTGGGTCTCAGCCGGAGCA 786  
||||| 902 CCATTTGACGTCGCTTATATTTGGGAAGTCTCTCTGTTTATTTCACTCACGCCGAGAGCAC 961  
||||| 787 CCACTTCTTTCCGTTATCGTCTCTACACTTCTTGTCTTCTTTCAGAGCGGAGCGTCAC 846  
||||| 962 GGTGTTGAAGTATGAGAGAACTGAAGAAGGTTTGAATCCACATCACTCACAATCTG 1021  
||||| 847 GGAATCAGCGTCACTCGGAAATTTACCAGAAAGGTTTGAATCCACCGTCTTGGAAATGCTT 906  
||||| 1022 GTATTTGTGTCGCTTACATGACTACAGCTGTCAAAACCTGCGATTTGCTTGGCATCAT 1081  
||||| 907 CAGTTTCACGGTAGTCACTTCGCACCTGTCGCCAAAACCGGACTCGTACCGGAATCGTC 966  
||||| 1082 TCACTTTCGCGAAGGAATGAGTAGGAAAGAGCTTTTGCATATGTATAAAAATTTACAATATT 1141  
||||| 967 TCCTTCACGGAAGGAATCGCAGTGGGAGAACATTTGCAGCGCTAAGAACTACCAGTA 1026  
||||| 1142 GATGGCAACAAAGAGATGATAGTATTTGGGACCATGAAGCTAGTGGTCTTTTCACTCT 1201  
||||| 1027 GATGGAAACAAAGAGATGATCGGCATTTGGTCTGATGAACCTAGTAGGCTCTGCGACATCT 1086  
||||| 1202 TGCTACTCTCAACAGACGACATTTTCGCGTTCGCTGTGAACCTATAACGTGATGCAAG 1261  
||||| 1087 TGCTAGTCTCAACCGGAGCAATTTCTTAGATCAGCGGTCAACAAACGCGGGAGCTAA 1146  
||||| 1262 ACAGCAGCTTCCAAACATTAATATGTCACTTTGCAAGTAATTTGACATTTGTTATTTCTTGACA 1321  
||||| 1147 ACCGAGTTTCAACATTTTATGTGCGTCACTGTTATGTTAGCTTCTCTCTTAATG 1206  
||||| 1322 CCCTTGTTCATTAACCTCCCTGGTGGTGTACAGCTATTAATCGTATCTGCAATGCTT 1381  
||||| 1207 CCGCTTTTCGAATACACTCCCAATTTGGTCTCGGTGCCATCATTTGTACCGCGGTCAIT 1266  
||||| 1382 GGACTCATAGATTTAGCAGCGCATCCATCTATTTAAGGTTGACAAATTTGACTTTGTG 1441  
||||| 1267 GGTCTCATCGACCTTCCCGGGCTGTGCATATGGAAGATCGATAAATTTGATTTCTTG 1326  
||||| 1442 GTGTGATGATGATACATACATTTGGTGGTCTTTGGCAGTGTTCGAAATTTGGCTTAGTCATA 1501  
||||| 1327 GTGATGCTTTGCGGGTCTTTGGTGTATTTTCTATCCGTTCAAAACGGTCTAGCCATT 1386  
||||| 1502 GCTATTGTAATATCTGTACTTCGGGTACTTCTATTTATGCAAGGCCAAGCAATTCGTT 1561  
||||| 1387 GCGGTGGGGCTATCGTTTGTTCGAAGATTTGATGCAAGTAAAGGCGGAAAATGGTTATA 1446  
||||| 1562 TTGGGCAACATTTCCAAATTTCTGTATATACCGAATGTTGAGCACTATCAAAATGCAAA 1621  
||||| 1447 ATGGGTAATATTTCTGGAACGGATATATACCGAGATCTTCATCATTTACAAAGAAGCACAA 1506  
||||| 1622 CATGTTCTCTGGAATGCTAAATTTCTAGAGATTTGATGACCAATTTTACTTTTGGCAATGCGAGC 1681  
||||| 1507 AGATCCCGGGATTTCTGTTTAAAGCATCGAATCTCTCTCAATTTTCGCAATTTCTAAC 1566  
||||| 1682 TATTTAAGAGAAAGGATCACAAGGTGATTTGATGAAGAAGAAAGAAATTTAAAGCTACA 1741  
||||| 1567 TACCTCACTGAAAGAACATCTCGTTGGATTGAAGATGCGAAGAGGAAAGCTCAAGAG 1626  
||||| 1742 GGGGAGACTAGTTGACGATGTTATAATTTGATATGAGTCTGTTGGAAACATTTGATACA 1801  
|||||

Db 1627 AAGCATTTAGCCTACAGTCTTGTGATCTTGAATGTCAGCCGTCAGCGGTAGACACA 1686  
Qy 1802 AGTGGAAATAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGCTACAGCTT 1861  
Db 1687 AACGGAGTGTCTTTTTTAAAGGAACTTAAGAAAAACAACGCCCAAGAGGACATCGAGCTT 1746  
Qy 1862 GTTTTGGTCAATCTCTAAGTGAAGTGAAGAACTGAAGAACTGAAGAACTGAAGAACTGAAGAACT 1921  
Db 1747 GTGTTTGTGAACCTCTTAAGCGAAGTGTGGAGAAGCTTTCAAAGAGCTGACGACAAAAA 1806  
Qy 1922 CATTTAGGGAAG--AAATGGATCTATCTGACTGTTGTAAGAGGCCGCTTG 1967  
Db 1807 GAGTTTCATGAGGCCCGAGTTTCTTCTTTAAACCTGCTGAGGCCGCTTG 1855

RESULT 11  
AAZ50481  
ID AAZ50481 standard; cDNA; 2279 BP.  
XX AAZ50481;  
AC AAZ50481;  
DT 23-MAY-2000 (first entry)  
XX  
DE Corn sulphate permease-1 cDNA clone.  
XX  
KW Sulphate Permease; sulphate assimilation protein; corn; contig; probe;  
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;  
KW antibody; screen; ss.  
XX  
OS Zea mays.  
XX  
FH Key  
CDS 3..2069  
FT /\*tag= a  
FT /product= "Corn sulphate permease-1"  
FT /notes= "Derived from a contig composed of clones  
FT cbn10.pk0062.b10, ccoin.pk081.b21, ccoin.pk092.12,  
FT csc1c.pk005.j3, p0004.cb1ej58r, p0089.csdch19r,  
FT p0094.csssg12r, p0121.cb1mx30r and p0128.cb1mx30r"  
FT  
XX  
PN WO200004154-A2.  
XX  
PD 27-JAN-2000.  
XX  
PF 13-JUL-1999; 99WO-US015810.  
XX  
PR 14-JUL-1998; 98US-0092833P.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
PI Allen SM, Falco SC, Thorpe CJ;  
XX  
XX WPI: 2000-195025/17.  
DR P-PSDB; AAY44935.  
XX  
XX Nucleic acid fragments encoding sulfate assimilation proteins in plants  
PT and seeds useful as probes for isolating cDNAs and genes encoding  
PT homologous proteins, in producing transgenic plants.  
XX  
PS Claim 3; Page 38-39; 79pp; English.  
XX  
CC The present sequence is a cDNA clone encoding corn sulphate permease, a  
CC sulphate assimilation protein. This sequence is from a contig composed of  
CC clones cbn10.pk0062.b10, ccoin.pk081.b21, ccoin.pk092.12, csc1c.pk005.j3,  
CC p0004.cb1ej58r, p0089.csdch19r, p0094.csssg12r, p0121.cb1mx30r and  
CC p0128.cb1mx30r from cbn10, ccoin, csc1c, p0004, p0089, p0121 and  
CC p0128 libraries, respectively. This sequence is used as a probe to  
CC isolate other plant sulphate assimilation proteins, for genetic and  
CC physical mapping of related genes and as markers of traits linked to the  
CC gene. This is useful for plant breeding and to construct chimeric genes,  
CC used to create transgenic plants with altered levels of sulphate  
CC permease. The sulphate permease peptides are useful for producing  
CC antibodies, that are used to screen and isolate cDNA clones



XX	Sequence	2279 BP; 512 A; 606 C; 629 G; 532 T; 0 U; 0 Other;	
SQ	Query Match	26.8%; Score 655.4; DB 3; Length 2279;	
	Best Local Similarity	60.4%; Pred. No. 3.8e-151;	
	Matches 1120; Conservative	0; Mismatches 726; Indels 9; Gaps 2;	
QY	121	AGTGCAACCAAGTCAGAGTTCACCGCCACAGACCGGTTTTCAAGTCTCTAAAGTACTCTTT	180
DB	188	AGTGCAACAGAGTGGCGCGCCACCGCGGCGGAGCACGGCGAGCAAGATGAAGTGAGG	247
QY	181	GAAGGAGACTTCTCCCTGANGACCTTTGAGGCAGTTCAAGAACAAAGCCAGCTTCCAA	240
DB	248	GAAGGAGACCTTCTTCCCGACGACCCGTTCCGGGCGTTCAAGGGGCGAGCGCGGGGAC	307
QY	241	GAAGTTATGCTTGGCTTTCAGTCTCTTCTCCCATTTTCGAATGGCTCCCAATACAC	300
DB	308	GCATGGCTCATGGCGGTTCAGGTACCTCTTCCCATCTCTGGACTGGGTCCGAGCTACTC	367
QY	301	CTTTCAAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCC	360
DB	368	CTTGTGCTCTTCAAGTCCGACCTCGTCGGGGCTTCCATTTGCCAGCTCGCCATTCC	427
QY	361	TCAGGCAATCAGTTATGCCAAGCTCGCCAACTTCCCTCCAAATCTTTGGACTATATTTCGAG	420
DB	428	TCAGGCAATTAGCTAGCGGAAGCTGGCAAGCTTGCCTCCCATTAATCGGGCTGTATTTCGAG	487
QY	421	CTTTATACCAACCATGATTTATTCGATGATGGTAGCTCGAGGGATTTGGCAGTGGGAC	480
DB	488	CTTGTGCGCCGATGGGTACGCGGTGTCTGGGGAGCTCCCGTACCTGGCGGTGGGCCC	547
QY	481	TGTGGCGGTTGATCCGCTTCTGATGGGTTCCGATTTGAGTAATGCCGTTGATCCCAATGA	540
DB	548	GGTGTGATCTCGTCTGCTGATCATGGGTCCATGCTCGGCAGCGCTGAGCCCCACTGC	607
QY	541	AGACCCAAAGCTTTACTCACCTGCTTTTCAAGCTATACATTAATTTGCTGGTGTTTTCA	600
DB	608	GGAGCCGAGCTGTCTCGCAGCTGGCTTCACTCCACCTGTTCCGGGGCTGGTGCA	667
QY	601	GGCTGCTTGGGTCGTTTGGTGGGTTGATCGTGGATTTCTGTCACATGCCAACAT	660
DB	668	GGCTTCCCTGGGATCCTCAGGCTCGGCTTCGTCATCGACTTCCGTCCAAAGCGACGCT	727
QY	661	AATAGGGTTTCATGGAGGAGCAGCACGCTGGTGTCTGTCAGCACTAAATTCGATTCT	720
DB	728	GGTGGGTTTCATGGCGCGCGCCCATCATCTGGGCGCTGCAGCACTCAAGGGCTGCT	787
QY	721	TGGCTTTGAGCATTTCCACCATGGAGCTGATATCATATCAGTGATGCGCTCTGTTTTCAC	780
DB	788	GGGCATCGTCCACTTCACCAACGAGATGGCATCGTCCCAGTCATGGCTCGTCTTTCCA	847
QY	781	CCAACTCATGATGGAGTGGGAAGTCTGTGTTAGGATGTGCTTCATTTCTTCTCT	840
DB	848	CCACACGAGGAGTGTGTGGGAGACGATCCCTCATGGGCGTCTGCTTCTCGTCTTCT	907
QY	841	CCTTAGCAAGAATCTTCAGCAAAAACGACCAAGGTTTTTGGGTGTGAGCAATGGC	900
DB	908	GCTGTGGCGGAGCATGTGAGCATCAGATGGCCAAAGCTTTCTGGGTTTCGGCGTGGC	967
QY	901	GCCATTGAGCTCGGTTATATGGGAAGTCTCTTGGTTTATTTTCACTCACGCCGAGAGCA	960
DB	968	GGCCCTGGCATCGGTCAACCATCTCGACGCTGTGTTTCTTCTTCCATCAAGCTCAGAACA	1027
QY	961	CGGTGTTGAAGTATAGGAGACTGAAGAAGGTTTGAATCCACCATCACTCACAAATCT	1020
DB	1028	TGGCATCAGCATCAITGGGAGCTCAAGTGGCGCTGGAATCGCCCTCGTGGGACAAGCT	1087
QY	1021	GGTATTGTGTGCGCTTACATCACTACAGCTGTCAAAACTGGCAATTTGCTGGCATCAT	1080
DB	1088	CCTGTTTGACCGCGTATTTAGGCTCTACCATGAGACTGGGCTGTTCACCGGAATCAT	1147
QY	1081	ATCATTGCGGAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTACAAAT	1140

DB	1148	CTCACTGACGGAAGGAATAGCGGTTGGTAGAAACATTTGCCTCACTCAAGGACTACCAGAT	1207
QY	1141	TCATGSCAAACAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTTTCACTTC	1200
DB	1208	AGATGAAACAGAGAGATGATGCCATAGGGTTGATGATGTTTGGGTCTTGCAATC	1267
QY	1201	TTGCTACCTTCACAAACAGGACCAATTTTCGGGTTTCGGCTGTGAACACTATAACCGCTGGATSCAA	1260
DB	1268	ATGCTACGTAAACACAGGTCGCTTCTCCGCTCTGCTGTAAACCAACACGCCGGCTGCAA	1327
QY	1261	GACAGAGCTTCCAAACATTAATATGCTACTTCGAGTAATGTTGACATTTGTTATTTCTGAC	1320
DB	1328	GACTGCCATGTCCAAAGCTGATCATGGCGCTGACTGTGATGGTCAGCTCTGTTCCTCAT	1387
QY	1321	ACCCTTTGTTCCATTACACTCCCTCGTGGTCTATCAGCTATTATCGTATCTGCAATGCT	1380
DB	1388	GCCACTGTTTGTGTACACCCCAAGCTTGTCTCGGAGCGATCATCATCGCCGGGTGAT	1447
QY	1381	TGGACTCATAGATTATGAAGCAGCCATCCATCTATTATTAAGTTTGAACAAATTTGACTTTGT	1440
DB	1448	CGGCTGATCGATTTCCCGCGGTGTACACATCTGGAAGATGCAACAGATGGAATTTCT	1507
QY	1441	GGTGTGATGATGATGATCAATTTGGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAAT	1500
DB	1508	GGTGTGCGTTTTCGGGTTTTCGGCGGTTCATCTTCTCAGTCCAAAGAGGCTTGGCAT	1567
QY	1501	AGCTATTGTAATATCTGTACTTTCGGGTACTTCTATTATTGCAAGGCCAAGACATTCGT	1560
DB	1568	ACGGTTGGTATATCTATATTTAGGGTGTGATGAGATCACAGGCCGGAAGATGATGT	1627
QY	1561	TTTGGGCAACATTTCCAAATTTCTGTATATACCGAAATTTTGAGCACTATCAAAATGCAAA	1620
DB	1628	TCAAGGGAACATCAAGGGGACTGATATTTACAGAGACCTGCATCACTCAAGAGAGCCCA	1687
QY	1621	ACATGTTCTGGAATGCTAAATTTAGAGATTTGATGACCAATTTACTTTGCCAATGCCAG	1680
DB	1688	AAGAGTTTCTGGGTTCTTGTATCTTGGCCATTGAAGCACCGATAAACTTCGCAACTCCAA	1747
QY	1681	CTATTTAAGAGAAAGATCACAAAGTGGATTCATGAAGAAAGAAAGAAATTAAGCTAC	1740
DB	1748	CTACCTGAATGAAGATTTAAAGATGGAT-----AGAGGAAGAAATCTTTTGAACAGGA	1801
QY	1741	AGGGGAGACTAGTTTGCAGTATGTTTATAATGATATGATGCTGTGTTGGAACATTTGATAC	1800
DB	1802	TAAACATCTGAACCTCCATTTTCAATCTTGTGATCTGTGAGTCTTCTGCAATTTGACAC	1861
QY	1801	AAGTGAATAAGTATGCTTGAAGAGGTGAAGAGATTTACAGAGAGAGAGAGCTACAGCT	1860
DB	1862	AAGTGGCATAGCGTTTCTCATTTGACATAAAGAAATCAATAGAGAAACGTTGGTCTGGAGCT	1921
QY	1861	TCGTTTGGTCAATCTGTAAAGTGAAGTGAAGAAACCTGAACAAATCGA---AGTTCCA	1917
DB	1922	TGTGCTGTGCAATCAACTGGAGAGTCAATGGAGAAATACAACTGCAACAGAGCTGA	1981
QY	1918	AAATCATTTAGGGAAGAAATGGAATCTATCTGATCTGTTGAAGAGCGGCTTTGGAGCA	1972
DB	1982	AAACTATTTAGGCCAGATTTGCTTGTATCTGACCACTGGCGAAGCAATCGCTTCA	2036

RESULT 12

ADA70466

ID ADA70466 standard; DNA; 1989 BP.

XX ADA70466;

XX AC ADA70466;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 3789.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice;

XX OS Oryza sativa.



Db 1669 TATGTTAAAGACAGATTCCTGAGGTGGCTGAGGACGAGGAAGACGTCAGCAGGAACAG 1728  
QY 1742 GGGGAGACTAGTTTGCAGTATGTTATAATTGATATGAGTCTCTTGGAACATTTGATACA 1801  
Db 1729 AAGCTACAGAAACTGAGTTTCTGATTTGTTGACTATCTCCGGTAATTGACATTTGATACA 1788  
QY 1802 AGTGGAAATAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTTACAGCTT 1861  
Db 1789 AGTGGAAATCCATGCTTGGAGGATCTATTTAGAGCACTTGAAAGCGCAAAATCCAGCTG 1848  
QY 1862 GTTTTGGTCAATCCTGTAAGTGAAGTGAAGAAGAACTGAACAAATCGAAGTTCCAAAT 1921  
Db 1849 ATTCTAGCCAATCCGGGGCAGCTGTGATCTCGAAGCTCCGATCCGATCAGCGAAATTCACGGAT 1908  
QY 1922 CATTTAGGGAAGAAATGGATCTATCTGACTGTGTTGAAGAGGCCGT 1965  
Db 1909 CTATCGGTGAAGACAAGATATTCCTGACTGTGCGGAGCGACT 1952

## RESULT 13

ADA69770  
ID ADA69770 standard; DNA; 1947 BP.

XX AC ADA69770;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 3093.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.

XX OS Oryza sativa.

XX PN WO2003000898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX DR WPI; 2003-175290/17.

XX PT Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.

XX PS Claim 6; SEQ ID NO 3093; 899pp; English.

XX CC The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.

XX SQ Sequence 1947 BP; 515 A; 481 C; 472 G; 479 T; 0 U; 0 Other;

Query Match 26.4%; Score 647.2; DB 8; Length 1947;

Best Local Similarity 60.3%; Pred. No. 3.8e-149;

Matches 1110; Conservative 0; Mismatches 718; Indels 12; Gaps 2;

QY 126 AACAGTCTGAGGTTCACCGCCACAGCCGCTTTTCAAGTCTCTAAAGTACTCTTTGAAGG 185  
Db 83 ACATGTCGGCGCGCGCGCGGAGAACTCTCTGCGGAGTTCGCGCGCAGGTGAAGG 142  
QY 186 AGACTTTCTTCCCTGANTGACCTTTTGAAGCAGTTCGAAGAACAGCCAGCTTCCAAGAGT 245  
Db 143 AGACGTTCTTCTCCAGCAGCGGATCGCGCGGTACAAGGACCGAGCGGTCCAGGAAGC 202  
QY 246 TCATGCTTGGCCTTCAGTTCCTTCTTCCCATTTTGGAAATGGGCTCCCAATACACCTTTC 305  
Db 203 TATGGCTGCCCTTGAGACACGCTTCTCCCGGTGTGGAATGGGGCAGACAATACACCTCG 262  
QY 306 AGTTCTTGAAGCTGACCTCATAGCTGCGCATCACCATCGCTAGCTTGGCCATTCCTCAG 365  
Db 263 CCAAGTTCAAGGGCGACCTCATTTGCGGCTCACCTTGGCAGCTCTGTCATACCTCAG 322  
QY 366 GCATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTGGAGCTATATTCGAGTTTA 425  
Db 323 ACATCGGCTACGGAAGCTTGCTAACTGCCACCAGAGATTGGGCTGCACAGTAGCTTCG 382  
QY 426 TACCACCATTTGATTTATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAGCTGCG 485  
Db 383 TCCCGCCATTGATATACGCTCTGATGGGCACCTCAAGGGAGCTAGCAATGGGTCCAGTGG 442  
QY 486 CGTTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAAATGCCGTTGATCCCAATGAAGACC 545  
Db 443 CGTCACTCACTGCTGCTTGGTACTCTCTCCAGGAGGAGATTGACTCAAGAGAACC 502  
QY 546 CAAAGCTTTACCTCCACCTGGCTTTCAAGCTACATTAATTTGCTGGTGTGTTTTCAGGCTG 605  
Db 503 CGCTAGATTACAGAGCGGCTCGCTTTCACAGCGACCTTCTTTGAGGAGTTCACAGCGCG 562  
QY 606 CCTTGGCTGCTTTAGTTGGGTTGATCGTGGATTTCTGTCACATGCAACCATTAATAG 665  
Db 563 CGCTGGTTTCTGAGGCTTAGGTTTCATATAGCGTTCTGTCATGCTGCATCATCG 622  
QY 666 GGTTCATGGGAGGAGCAGCAGCGTGGTGTCTGCAGCAACTAAATCGATTTCTTGGCC 725  
Db 623 GATTCAATGCGCGCGCGCCATCACATTTGCTCTTCCAGCAGCTTAAAGGCTTCTTTGGAA 682  
QY 726 TTGAGCATTTACCCATGGAGCTGATATCATATCAGTATGCGCTGTGTTTTCACCCAAA 785  
Db 683 TTGCAAACTTCAACCAAGAGACTGACATCATCTCGTTCATGAATCAGTCTGGGGAATG 742  
QY 786 CTCATGAGTGGAGGTGGGAAGTGTGTTTAGGATGTCTTCATTTTCTTCTCTCTTA 845  
Db 743 TTCACATGG-----GGCAAGTAATTTTGTACCTCTCCAACTTCTTGAT-CTCT 790  
QY 846 GCACAAGATACTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTCAAGATGGGCCAT 905  
Db 791 TCTCTGAAAAAGGGGCGAAGAGAACAAGAGCTCTTCTGGGTCCAGCAATTCACCCAC 850  
QY 906 TGACGTCGTTATATTTGGGAAGTCTCTTGGTTTATTTTCACTCAGCGGAGAGCAGCGTG 965  
Db 851 TCATTTCCGTGTGATCATTTCAACTTTTGTCTGTACATCATCTGCTGTGCAACAAGGCG 910  
QY 966 TTGAAGTGATAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCAAAATCTGGTAT 1025  
Db 911 TCGCAATCGTAAAAAAGCTCAAGAAAGCATCAATCCACCTTCAGTAGCCCTGATATTTT 970  
QY 1026 TTGTGTCGCTTTACATGACTACAGCTGTCAAACTGGCATTTGCTTGGCATATATCAC 1085  
Db 971 TCACTGGCCCATACTTGTCTCAAGGATTCAAAATTTGGAGTAGTAGTGAATGATAAGCC 1030  
QY 1086 TTGCGGAAGGATACAGTAGGAAGAAGCTTTGCAATGTATAAAAATTACAATATTGATG 1145  
Db 1031 TTACGGAAGCGATTTCAGTTGGAAGAACAATTTGCTGGATTGAACGATTAACAGATAG 1090  
QY 1146 GCAACAAAAGAGATGATAGCTATTGGGACCATGAACTGAGTTGCTTCTTTCACCTCTTGT 1205  
Db 1091 GGAAACAAGAAATGTTGGCTCTAGGAACCATGAATGTGGTTCGTTCAATGAGCTCTTGT 1150  
QY 1206 ACCTCACAAACAGGACCAATTTTCGCGTTCGGTGTGAACTATTAACCGTGGATGCAAGACAG 1265

Db 1151 ATATAGCCACAGGTGGTTTTCACGATCAGACGTCATATTCGATGGCTGGAGGTAAACAC 1210  
Qy 1266 CAGCTTCCAAACATTAATGTCTACCTTCAGTAATGTTGACATTTGTTATTCCTGACACCCCT 1325  
Db 1211 CAATGTCCAATATGTTATGTCAACTGTAGTATTTGCTTGACCTCTGTGGATCACTCCAT 1270  
Qy 1326 TGTTCATATACATCCCTCGTGTGTATCAGCTATATTCGTAATCTGCAATCGTTGGAC 1385  
Db 1271 TGTTCAGATACAGCCCAATGCCACCAATTTCTCCATCATCATATCAGCATGCTTTGGCC 1330  
Qy 1386 TCATAGATATGAAGCAGCCATCATCTATTTAAGGTTGACAAATTTGACTTTGTGTGT 1445  
Db 1331 TATTTGACTTTGATCAGCTCACTTATCTGGAAGTTGATAGTTGGACTTCATGCGGT 1390  
Qy 1446 GCATGATGCATACATTTGGCGTGTCTTTGGCGATGTTGAAATTTGGCTTAGTCATAGCTA 1505  
Db 1391 GCTTGGGGGATTCCTTTGGAGTAATAATTTTCATCTGTGGAGTATGGCTTGCTCATTTGCGG 1450  
Qy 1506 TTGTAATATCTGACTCTCGGCTACTTCTATTTATTCGAGGCGCAAGACATTCGTTTGG 1565  
Db 1451 TTGTAATATCACTAATCAAAAGTTCTGCTCCATGTAAACACGCGCAAGCAGCTTTTACTTG 1510  
Qy 1566 GCAACATTTCCAAATTCGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACATG 1625  
Db 1511 GCAACCTTCCAGAACGATTTATCTATAGAAATGTTGAACAATATCCAGAACTACCAAGG 1570  
Qy 1626 TTCTGGAATGCTAAATCTAGAGATGATGCACCAATTTTACTTTGCAATGCCAGCTATT 1685  
Db 1571 TGCAGGAGTCTAATTTGAAGTGGACTCAGCAATATATCTTCACAAACTCCAAATATG 1630  
Qy 1686 TAGAGAAAGGATCACAAGGTGATGATGATGATGATGATGATGATGATGATGATGATG 1745  
Db 1631 TTAAGAAAGAAATGCTGAGATGCTGAGATGCTGAGATGAGGAGAAACATCAAAAGGAAACAG 1690  
Qy 1746 AGACTAGTTTGCAGTATGTTATATATGATATGATGATGATGATGATGATGATGATGATG 1805  
Db 1691 TACCAAAATGATTTCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1750  
Qy 1806 GAATAAGTATGCTTGAAGAGGTGAAGAAGATTAACAGAGAGAGAGAGATACAGCTTTGTT 1865  
Db 1751 GAATCCATGCTTCAAAAGAGTGTGAGGACACTTGAAGAGCGCCAGATTCAGCTGATTT 1810  
Qy 1866 TGCTCAATCTGTAAAGTGAAGTGAAGAACTGAACAAATGAAGTTCGAAATCAAT 1925  
Db 1811 TCGCCAATCTCGGGCGGCTGTGATCCAAAGCTCCGGTCAGCAAAATTCACAGAGCTCA 1870  
Qy 1926 TAGGGAAGAAATGATCTATCTACCTGTTGAAGAGCGCT 1965  
Db 1871 TTGGTGAAGAAAGATATGCTGACAGTTGGTGACGCCGT 1910

RESULT 14

AAZ50490

ID AAZ50490 standard; cDNA; 2311 BP.

XX AC AAZ50490;

XX DT 23-MAY-2000 (first entry)

XX DE wheat sulphate permease-1 cDNA clone.

XX KW Sulphate Permease; sulphate assimilation protein; wheat; probe; mapping;

XX KW marker; plant breeding; chimeric gene; transgenic plant; antibody;

XX KW screen; ss.

XX OS Triticum aestivum.

XX Key Location/Qualifiers

XX CDS 100..2070

XX FT /\*tag= a

XX FT /product= "wheat sulphate permease-1"

XX FT /note= "Derived from clone wki.pk0028.el"

XX PN WO200004154-A2.  
XX PD 27-JAN-2000.  
XX PF 13-JUL-1999; 99WO-US015810.  
XX PR 14-JUL-1998; 98US-0092833P.  
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX PI Allen SM, Falco SC, Thorpe CJ;  
XX WPI; 2000-195025/17.  
XX P-PSDB; AAY4944.  
XX Nucleic acid fragments encoding sulfate assimilation proteins in plants  
XX and seeds useful as probes for isolating cDNAs and genes encoding  
XX homologous proteins, in producing transgenic plants.  
XX Claim 3; Page 55-56; 79pp; English.  
XX The present sequence is a cDNA clone encoding wheat sulphate permease, a  
XX sulphate assimilation protein. This sequence is obtained from wki  
XX library, clone wki.pk0028.el, derived from wheat seedlings, 1 hour after  
XX treatment with fungicide. This sequence is used as a probe to isolate  
XX other plant sulphate assimilation proteins, for genetic and physical  
XX mapping of related genes and as markers of traits linked to the gene.  
XX This is useful for plant breeding and to construct chimeric genes, used  
XX to create transgenic plants with altered levels of sulphate permease. The  
XX sulphate permease peptides are useful for producing antibodies, that are  
XX used to screen and isolate cDNA clones  
XX SQ Sequence 2311 BP; 670 A; 493 C; 537 G; 611 T; 0 U; 0 Other;

Query Match 26.1%; Score 640.4; DB 3; Length 2311;

Best Local Similarity 59.4%; Pred. No. 1.9e-147; Mismatches 751; Indels 3; Gaps 1;

Matches 1104; Conservative 0;

Qy 122 GTGCACCAAGTCGAGGTTCCACGCCACACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTG 181  
Db 187 GTGTACAAAGTGGGCTATCCCTCCGAGAACTTGGCCACAGAGTTTACAGAAACATTG 246  
Qy 182 AAGGAGACTTTCTTCTTCCTGATGACCTTTTGAGGAGCTTCAAGAAACAGCCAGCTTCCAAG 241  
Db 247 AGGAGACTTTCTTCCACGCAACCCGCTGCTGCTAGTATAAGGGCCAATCCGACCGCAGG 306  
Qy 242 AAGTTCACTCTTGGCTTCAGTTCTTCTTCCCATTTTTCGAATGGCTCCCAATACACC 301  
Db 307 AGGTTCACTGATGGGCTGGAGTTCTTGTTCCTATATTTGGGTGGGTAGGGATTACAGT 366  
Qy 302 TTTTCAGTTCTTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCTCT 361  
Db 367 CTCACAAAGTTCAAGGCGCATCTGATGCGGAGTGGACCATTCGCAAGTCTCTGTATTCTCT 426  
Qy 362 CAGGGCATCAGTTATGCCAAGCTCGCAACCTCCCTCCAATTTCTTGGAATAATTTCGAGC 421  
Db 427 CAGGACATTTGGCTATTGGAAGCTTGCTAATCTGGATCCGAGTATGGGCTTTACTCCAGC 486  
Qy 422 TTTATACCAACCATTTGATTTTATCGGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACT 481  
Db 487 TTTCAATCTCTCCATTTGATCTATGCTGCAATGGGTAGCTCAAGGATATAGCGATTGGTCCA 546  
Qy 482 GTGGCGTTGGATCGCTTCTGATGGGTTCGATGTTTGAAGTAAATGCCCTTCGATCCCAATGAA 541  
Db 547 GTTGTGTGGTTTCTCTTTTGTAGTTTCACTTCTTCAAGCTGAGGTTGACCATGTCAAA 606  
Qy 542 GACCCAAAGCTTTTACCTCCACCTGGCTTTCACAGCTTACATTTATTTGCTGGTGTTCAG 601  
Db 607 AACAGGAGGAATACATGCGCTCGCTTTTTCACGGCAACCTTCTTCTGCTGATCACTCAA 666  
Qy 602 GCTGCTTGGGTCTGTTTAGTTGGGTGATCGTGGATTTTCTGTCAATGCAACCATTA 661

Db 667 GCAGCCTTAGGATTTCTAAGTTAGGATTCCTTATAGAGTTCTTGTGCGATGTCGGATT 726  
QY 662 ATAGGGTTCAATGAGGAGCAGCCAGGTGCTGTCTGCAGCAACTAAATCGATTCTT 721  
Db 727 GTCGGATTCATGGGGAGCTGCCATTACTATTGGCCCTGCAGCAGCTGAAATACGTGTG 786  
QY 722 GGCCTTTGAGCATTTTCAACCCATGAGCTGATATCATATCAGTGTGCGCTCTGTCTTTCACC 781  
Db 787 GGCATCGCAAACTTTACAAGGAAACCGACATAGTTTCTGTCAAGAACTGTCTGAGAA 846  
QY 782 CAACATCAT---GAGTGGAGGTGGAAAGTGTGTTAGTAGTGTCTTCAATTTCTTC 838  
Db 847 TCAGTTTCATACGGGTGGAACTGCGACAAATTTGATTGGCGTATCTTTTCTGGTTTC 906  
QY 839 CTCCTTTAGCACAAGATACTTTCACAAAAACGACCAAGGTTTTTTTGGGTGTGACCAATG 898  
Db 907 CTTCTGTTTGGGAAGTACATCGGAAGAGAAAGAAAGCTTTTCTGGGTGCCAGCTATT 966  
QY 899 GCGCCATTGACGTCGCTTATATTGGGAAGTCTCTTGGTTTATTTCCTCAACCGGAGAA 958  
Db 967 GCTCCTATAATTTTCAGTGATTTAGCAACATTTTGTATACATTACTGTCGCGACAAG 1026  
QY 959 CACGGTGTGAAGTAGAGAACTGGAAGAGGTTTGAATCCACCATCACTCACAAT 1018  
Db 1027 CAAGGAGTTCAGATAGTGAAGCAATTGAACAGGGAATCAACCCATCATCAGTACACAAG 1086  
QY 1019 CTGGTATTGTGTCGCTTACATGACTACAGCTGTCAAACTGGCAATTCGTTGGGCATC 1078  
Db 1087 ATTATTTTACCGGCCAATTTGTGCAAGAGTTTCAAGATCGGTGTTTGTGGCGCAT 1146  
QY 1079 ATATCACTTTCGGAAGGAATAGCAGTAGGAAGCTTTGCAATGTATAAAATTTACAAT 1138  
Db 1147 GTTGGTTTGACAGAAGCTGTAGCTATTGGAAGGACATTTGCTGTATGAAGACTACCA 1206  
QY 1139 ATTGATGGCAACAAAGAGATGATAGCTATTGGGACCAATGAAGTGTGGTTCTTTTTCAC 1198  
Db 1207 TTAGATGGAACAAAGGAGATGGTAGCACTTGGAAACCAATGAATAGTGGCTCAATGCA 1266  
QY 1199 TCTTGCTACTCACAACAGACCATTTTCGCGTTCCGCTGCGCTGTGAATTAACGCTGATGC 1258  
Db 1267 TCTTGCTATGTCAACAGGTTCTTTCTCACGTTCCGCGAGTTAACTTTCATGGCTGGCTGC 1326  
QY 1259 AAGACAGCAGCTTCCAAACATTAATAATGTACATTTGCAGTAATGTTCATATTGTTATTTCC 1318  
Db 1327 AAGACTCCTGTATCCAAATGTGGTTATGTAGTAGTGTCTTCTTACCTTGTGTGTCATC 1386  
QY 1319 ACACCTTGTTCATTAATACACTCCCTGGTGGTGTGCTATCAGCTATTATCGTATCTGCAATG 1378  
Db 1387 ACACCGCTATTCAAATATATACCCGAATGCAATCCTAGGGTCGATCATATTATTTCTGCGGTG 1446  
QY 1379 CTTGGACTCATAGATTATGAAGCAGCCATCCATCTATTATTAAGTTTGACAAATTTGCACTTT 1438  
Db 1447 ATCGGCTTGTGGACTACGAAGCAGCAATTCATCTGGAAGTTGACAAATTTGCACTTC 1506  
QY 1439 GTGGTGTGATGATGATACATATGGCGTGGTCTTTGGCAGTGTGTGAAATTTGGCTTAGTC 1498  
Db 1507 ATTGCTTGATGAGGAGCTTTTTCGGGTGTTGTTTGTATCGCTTGAGATTGGCCTCTTG 1566  
QY 1499 ATAGCTATTGTAATATCTGTACTTCGGGTACTTCTATTATTATGCAAGGCCAAGGACATTC 1558  
Db 1567 ATTGCTGTAGCAATCTCAATTTGCCAAATTAATCTTCTTCAAGTAACCAAGGCCAAGGAC 1626  
QY 1559 GTTTTGGGCAACATTCCAAATTCGTGATATACCGAAATGTTGAGCACTATCAAAATGCA 1618  
Db 1627 CTACTTTGGAACTTCCGGCACCATATATACCGGAACATCAGCCAGTATCCAGAAAGCA 1686  
QY 1619 AAACATGTTTCTGGAAATGCTAAATCTAGAGATTGATGCAACCAATTTTACTTTTCCCAATGCC 1678  
Db 1687 AAACCTTACTCTGGGGTGTGATTGTGAGGGTTGATTCTGCTATTATTATTTTCCAACTCT 1746  
QY 1679 AGCTATTTAAGAGAAAGGATCAACAGTGGATTGATGAAGAGAAAGAAATTAAGCT 1738  
Db 1747 AATTACGTCGGAAGAAAGATTTCTTAGTGGCTGACAGACGAAAGACAGCAGCTAAAGCA 1806

QY 1739 ACAGGGGAGACTAGTTTGCAGTATGTTATAATTGATATGAGTCTCTTTCGAAACATTGAT 1798  
Db 1807 GTGGGATTCCCTAAATCAATGTTTCCGTGATTGTGGAATGTCGCCGTCTATCGACATCGAT 1866  
QY 1799 ACAAGTGAATTAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAAGAGAGCTACAG 1858  
Db 1867 ACAAGCGGCATACGCTCTTGAAGATCTATACAAGAACTTTTCAGAAAAAAGATATGAG 1926  
QY 1859 CTTGTTTGGTCAATCTCTGTAAGTGAAGTGAAGAACTGAAACAATTCGAAGTTCCTCAA 1918  
Db 1927 CTCATTCTGTCAATCTCTGGTTCCGTCGTCATAGAAAACTGCAAGCGTCGAAGCTCACC 1986  
QY 1919 AATCATTTTAGGGAAGAAATGGATCTATCTGACTGTGTAAGAGGCGCTTTCGAGCATGCA 1976  
Db 1987 GAGCATTGGGAAGCAGCAATATATCTTCGCGGTCTCTGACGCTGTGCGATTCTGTA 2044

RESULT 15  
ABQ82719  
ID ABQ82719 standard; cDNA; 1971 BP.  
XX ABQ82719;  
AC ABQ82719;  
DT 03-JAN-2003 (first entry)  
XX Arabidopsis thaliana sulphate transporter Sultr1;3 encoding cDNA SEQ:1.  
XX Arabidopsis thaliana; sulphate transporter; chromosome 1; plant;  
KW sulphur-containing metabolite; glutathione; phytokeatin; gene; ss.  
XX Arabidopsis thaliana.  
XX Key Location/Qualifiers  
FT CDS 1..1971  
FT /\*tag= a  
FT /product= "sulphate transporter Sultr1;3"  
XX JP2002272472-A.  
XX 24-SEP-2002.  
XX 22-MAR-2001; 2001JP-00082891.  
XX 22-MAR-2001; 2001JP-00082891.  
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
XX WPI; 2002-718707/78.  
XX P-PSDB; ABP53777.  
XX A gene encoding sulfate ion transporter Sultr1;3, a polynucleotide, an oligonucleotide, a recombinant vector, a plant.  
XX Claim 2; Page 6-9; 12pp; Japanese.  
XX The present invention describes a gene which is present in the first chromosome of Arabidopsis thaliana and encodes sulphate ion transporter Sultr1;3 having the amino acid sequence given in ABP53777. Also described: a polynucleotide purified from the genomic DNA, the mRNA, the cDNA or their complementary sequences; an oligonucleotide hybridising with the above gene or the above polynucleotide and consisting of a base sequence of at least 10 base pairs; a recombinant vector carrying the above polynucleotide; and a plant in which the above gene is introduced or its descendant or their tissues. The plant can be used for the preparation of sulphur-containing metabolites such as glutathione and phytokeatin. The present sequence encodes Arabidopsis thaliana sulphur transporter Sultr1;3, from the present invention  
XX Sequence 1971 BP; 504 A; 465 C; 427 G; 575 T; 0 U; 0 Other;  
Query Match 25.9%; Score 634.4; DB 6; Length 1971;  
Best Local Similarity 59.1%; Pred. No. 5.5e-146;

Matches	1104;	Conservative	0;	Mismatches	761;	Indels	3;	Gaps	1;
Qy	122	GTGCACCAAGTCGAGGTTCCACGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTG	181						
Db	88	GTCCACAAAGTCGAAGTTCCTCCTAAGCAAAACCTTTTCAATGAGTTCATGTACACTTTT	147						
Qy	182	AAGGAGACTTTCTTCCCTGATGACCCCTTTGAGGCGAGTTTCAAGAACCAAGCAGCTTCCAAG	241						
Db	148	AAAGAAACTTTTCTTCCAAGATGATCTCTAAGGCATTTTCAAGGACCAAGTCAAAATCCAAA	207						
Qy	242	AAGTTCAATGCTTGCCTTCTAGTTTCTTCCCTCCCAATTTTCCAAATGGGCTCCCAATACACC	301						
Db	208	AAGCTCAATGCTCGGTATCCAGTCCGTCTTCCGGTTATCGAGTGGGGAAGAAATATAAT	267						
Qy	302	TTTCAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTTGGGCATTCCT	361						
Db	268	CTTAAAGTTGTTTCGGCGCATCTTATTGCGGTTTAAACATAGCCAGTCTCTGCATTTCT	327						
Qy	362	CAGGGCATCAGTTATGCGCAAGCTCGCCAAACCTCCCTCCAAATCTTGAGACTATATTCGAGC	421						
Db	328	CAGGATATTGGATATGCAAGACTCGCGAGTCTTGAACCTAAGTATGGTCTATATTCAAGT	387						
Qy	422	TTTATACCCACATTGATTATTCGATGCGTAGCTCGAGGGATTTGGCAGTGGGGACT	481						
Db	388	TTTGTCTCCGTTGGTGTACGATGATGGGAAGCTCAAGGATATAGCGATTGGACCC	447						
Qy	482	GTGCGGTTGGATCGCTTCTCATGGGTTCCAGTTGAGTAATCCGTTGATCCCAATGAA	541						
Db	448	GTTCAGTGGTTTCACTCCTATTAGGTACTCTGCTCGTCTGAGTCGACCCCAACACA	507						
Qy	542	GACCCAAAGCTTTACCTCCACCTGGCTTTCACAGCTACATATTGCTGTGTTTTTCAG	601						
Db	508	AACCCATAAGTAATCTCCGTTAGCCTTCACTCCACCTCGAGTTCTTTGCGGTGTCACTCA	567						
Qy	602	GCTGCCTTGGTCTGTTTAGTGGGTGCGTGTGAGTATTTCTGTACATGCAACCATTA	661						
Db	568	GCAGCACTCGGATTTCTTCAGATTGGGTTTCTTGATCGAATTTCTTGTCGCCGCGGGTG	627						
Qy	662	ATAGGGTTTCATGGAGGAGCAGCCAGCGTGTGCTCGCAGCAACTTAAATCGATTCTT	721						
Db	628	GTAGTTTCATGGCGGAGCAGCCATCACATTCGCGTGCACAAGTCAAGGCTTCCTA	687						
Qy	722	GGCCTTGAGCAATTTCAACCATGGAGCTGATATCATATCAGTGTGCGCTCTGTTTTCAACC	781						
Db	688	GGAATCAATAAGTTTCAAAAGAAAACCGATATCATCGCGTCTTTCTTCGTAATCAGC	747						
Qy	782	CAAAC---TCATGAGTGGAGGTGGAAAGTGTGTGTAGGATGTGCTTCATTTTCTTC	838						
Db	748	TCAGCCCATCACGGATGGAAATTCGCGACAAATACTCATAGTCATCGTTTCTTGATCTTC	807						
Qy	839	CTCCTTAGCACAAAGATACTTTCAGCAAAAACGACCAAGGTTTTTTGGGTGTCAGCAATG	898						
Db	808	CTTCTCATCTCCAAGTTTATCGGAGAGAGAAACAAGAACTGTTTTGGATTCAGACTATT	867						
Qy	899	GGCCCATTCAGCTCGTTATATTGGGAAGTCTCTTGGTTTTTATTTCATCACGCCGAGAAG	958						
Db	868	GCTCCGTTAGTATCTGTGCATCATTTCAACCTTCTTCGTCTACATAACCCGAGCCGACAAG	927						
Qy	959	CACGGTGTGAAGTATAGGAGAACTGAAGAGGGTTTGAATCCACATCACTCAACAAT	1018						
Db	928	AAAGGAGTTTCAAGTATGAGAAACATCTTTGAACAAAGGTCTGAAACCTCTCTTTGCGGTCTA	987						
Qy	1019	CTGGTATTTGTGCGCTTACATGACTACAGCTGTCAAACTGGCATTTGTTGGCATC	1078						
Db	988	ATATATTTCTCGGCGAATTAACCTTCTCAAGGCTTTCGCAATAGCGTTGTCTCAGGCATG	1047						
Qy	1079	ATATCACTTCGGNAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATAAAAAATTACAAT	1138						
Db	1048	GTTCGCTTTCAGCGNAGCTGTAGCGATAGGAAGAACATTTTCGACCAATGAAGACTACCAA	1107						
Qy	1139	ATTGATGGCAACAAAGATGATAGCTATTGGGACCAATGAACGCTAGTTGGTTCTTTCAACC	1198						
Db	1108	ATCGATGGTAAACAAGAGATGGTAGCATTAGGAGCAATGAACGTAACTCGGTTCAATGACC	1167						

Search completed: August 31, 2005, 03:36:07  
Job time : 1320 secs

Qy	1199	TCTTGCTACCTCAACAACAGACCATTTTCGGTTCGGCTGTGAACTATAACGTCGGATGC	1251
Db	1168	TCTTGCTATGTATCCACCGGTCTTTCTCAAGATCCGCCGTCAACTTTATGGCGCGATGT	1227
Qy	1259	AAGCAGCAGCTCCCAACATTTAATGTCACTTCGAGTAATGTTGACATGTTATTCCTG	1318
Db	1228	CAAACGGCAGTCTCCAACATCATCATGTCAATTTGTGTCTCTTAACGCCTTCTTTCCTC	1287
Qy	1319	ACACCCCTGTTCCATTACACTCCCCCTGGTGGTCTATCAGCTATTAATFCGTATCTGCAATG	1378
Db	1288	ACTCCTCTTTTCAAAATACACACCCACGCAATTCCTTGCACGGATCATCATCAACGCTGTG	1347
Qy	1379	CTTGGACTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTT	1438
Db	1348	ATTCCTTTGGTGTACCGTTAATGTACTACATTTTGATCTTCAAGATCGAATAAGTCGATTTT	1407
Qy	1439	GTGGTGTGCATGAGTGCATACATTGGCGTGGTCTTTGGCAGTGTGAAATGGCTTAGTC	1498
Db	1408	GTTGCTGTATGGGGCTTTTTTGGTGTCACTCTTTGTATCCGTAGAGATTGGCTTCTA	1467
Qy	1499	ATAGCTATTGTATAATCTGTACTTTCGGGTACTCTCTATTATTATGGAAGGCCAAGGACATTC	1558
Db	1468	ATAGCCGTGGGGATATCTTTTGGCAAGATACITTTTGCAGTTTACGAGACCTAGGACAGCG	1527
Qy	1559	GTTTTGGGCAACATTCCAAATCTGTGATATACGGAAATGCTTGACGACTATCAAAATGCA	1618
Db	1528	ATTCCTTGGAAAGATACCGGGACTTCGGTTTTACAGGAATATCAATCAGTATCTGAAAGCG	1587
Qy	1619	AAACATGTTCTCGAATGCTAA'TTCTAGAGATTGATGCACCAATTTACTTTGGCAATGCC	1678
Db	1588	ACTAGGATTCGGGAGTTTTTGACAA'TTCGTGTTGACTCCGCGATTTACTTCTCCAATCC	1647
Qy	1679	AGCTATTTAAGAAAGGATCAACAGGTGATTTGATGAAGAGAAGAAATTAAGCT	1738
Db	1648	AATTATGTTAGGGAAAGGATTCAAAGATGGTTGACAGATGAAGAAGAGATGCTGGAGCT	1707
Qy	1739	ACAGGGGAGACTAGTTTGCAGTATGTTTATAATTTGATATGATGCTGTTTGGAAACATTGAT	1798
Db	1708	GCAAGATTGCTTAGGATCCAGTTTCTTATCATCGAAATGTCACCTGTTACGGACATCAT	1767
Qy	1799	ACAAGTGAATAAGTATGCTTGAAAGAGGTGAAGAATTAACAGAGAAGAGAGCTACAG	1858
Db	1768	ACTAGTGGTATTCAACGCCTTAGAAGACTTGATAAAGTCTCTCCAAAAACGAGACATTCAG	1827
Qy	1859	CTTGT'TTGGTCAATCCTGTAAGTGAGTATGAAGAACTGAACAAATCGAAGTTCCAA	1918
Db	1828	TTGGT'TCTACGGAATCCAGGACCGCGGTTCATAAATAAGCTACATGTTTCTCACTTTCG	1887
Qy	1919	AATCATTTAGGGAAAGAAATGGATCTATCTGACTGTTTGAAGAGCCGCTTGAGGATGCAAC	1978
Db	1888	GACTTGATAGGACACGCAAAATCTTCTGACCGTGGCTGAGGCCGTGGATTCTTGTCTCC	1947
Qy	1979	TTCAATCT 1986	
Db	1948	CCTAAACT 1955	

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 02:38:26 ; Search time 429 Seconds  
(without alignments)  
9340.890 Million cell updates/sec

Title: US-10-762-049-17  
Perfect score: 2449  
Sequence: 1 GCACGAGCTAGTGCACAT.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Databases : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/PCUTUS\_COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2449	100.0	2449	4	US-09-720-317A-17
2	765.2	31.2	1981	4	US-09-720-317A-3
3	720.6	29.4	2087	4	US-09-720-317A-15
4	655.4	26.8	2279	4	US-09-720-317A-1
5	640.4	26.1	2311	4	US-09-720-317A-19
6	324	13.2	780	4	US-09-720-317A-7
7	216.2	8.8	1240	4	US-09-720-317A-5
8	205.6	8.4	2022	4	US-09-720-317A-21
9	117.4	4.8	493	4	US-09-720-317A-13
10	117.4	4.8	484	4	US-09-720-317A-9
11	90.4	3.7	2441	4	US-09-785-381-4
12	82.6	3.4	4927	4	US-09-949-016-5627
13	79.8	3.3	4113	4	US-09-785-381-2
14	77.6	3.2	2487	4	US-09-248-796A-6643
15	76.4	3.1	7218	1	US-08-232-463-14
16	64.4	2.6	2682	4	US-09-614-221A-222
17	63.8	2.6	2913	4	US-09-795-927-6
18	63.8	2.6	3749	4	US-09-795-927-8
19	58.2	2.4	510	4	US-09-720-317A-11
20	58	2.4	758	4	US-09-270-767-470
21	58	2.4	758	4	US-09-270-767-15752
22	57.8	2.4	3061	1	US-09-570-842-1
23	57.8	2.4	3061	1	US-08-700-576-1
24	56	2.3	1773	4	US-09-902-540-7746
25	56	2.3	4854	4	US-09-902-540-768
26	53.4	2.2	2832	4	US-09-949-016-4516
27	53.4	2.2	8774	4	US-09-949-016-16258

28	53.2	2.2	320	4	US-09-270-767-27504	Sequence 27504, A
29	53.2	2.2	904	4	US-09-270-767-11855	Sequence 11855, A
c	52.8	2.2	1116	4	US-09-252-991A-7460	Sequence 7460, Ap
31	52.8	2.2	1791	4	US-09-252-991A-7241	Sequence 7241, Ap
32	51	2.1	761	4	US-09-270-767-1826	Sequence 1826, Ap
33	51	2.1	761	4	US-09-270-767-17108	Sequence 17108, A
34	51	2.1	1767	4	US-09-252-991A-7184	Sequence 7184, Ap
35	47.4	1.9	1971	4	US-09-875-811-9	Sequence 9, Appli
36	47.4	1.9	1992	4	US-09-875-811-5	Sequence 5, Appli
37	47.4	1.9	2040	4	US-09-875-811-1	Sequence 1, Appli
38	47.4	1.9	2420	4	US-08-232-463-14	Sequence 14, Appl
39	47.4	1.9	7218	1	US-08-232-463-14	Sequence 14, Appl
c	45.4	1.9	1141	4	US-09-806-708B-22	Sequence 22, Appl
41	45.4	1.9	1143	4	US-09-540-236-913	Sequence 913, App
c	45.4	1.9	31147	4	US-09-596-002-25	Sequence 25, Appl
43	45.2	1.8	1466	4	US-09-620-312D-914	Sequence 914, App
44	45.2	1.8	2858	4	US-09-949-016-5349	Sequence 5349, Ap
45	45.2	1.8	2882	1	US-08-424-567-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-720-317A-17  
; Sequence 17, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 17  
; LENGTH: 2449  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-720-317A-17

Query Match	100.0%;	Score 2449;	DB 4;	Length 2449;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2449;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	GCACGAGCTAGCTCGCACATTAAGTTATATACACATATTTGCTTCTTAGAAATACTAT	60	
DB	1	GCACGAGCTAGCTCGCACATTAAGTTATATACACATATTTGCTTCTTAGAAATACTAT	60	
QY	61	TATTGAAGATATGGGAGTGATGATTATGAGTACCTTTGGGCATCAACAACTTCGAGAG	120	
DB	61	TATTGAAGATATGGGAGTGATGATTATGAGTACCTTTGGGCATCAACAACTTCGAGAG	120	
QY	121	AGTGACCAAGTCGAGGTTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTT	180	
DB	121	AGTGACCAAGTCGAGGTTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTT	180	
QY	181	GAAGGAGACTTCTTCCCTGTGATGACCCCTTTGAGGAGAGTTCAAGAAACAGCAGCTTCCAA	240	
DB	181	GAAGGAGACTTCTTCCCTGTGATGACCCCTTTGAGGAGAGTTCAAGAAACAGCAGCTTCCAA	240	
QY	241	GAAGTTCATGCTTGGCTTTCAGTTCTTCTTCCCATTTTTCGAATGGCTCCCAAAATACAC	300	
DB	241	GAAGTTCATGCTTGGCTTTCAGTTCTTCTTCCCATTTTTCGAATGGCTCCCAAAATACAC	300	
QY	301	CTTTTCAGTTCTTGAAGAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCC	360	
DB	301	CTTTTCAGTTCTTGAAGAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCC	360	

QY	361	TCAGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAACTCTTGGACATATTCGAG	420
DB	361	TCAGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAACTCTTGGACATATTCGAG	420
QY	421	CTTTATACCAACCATTTGATTCGATGATGGGTAGCTCAGGGATTTGGCAGTGGGAC	480
DB	421	CTTTATACCAACCATTTGATTCGATGATGGGTAGCTCAGGGATTTGGCAGTGGGAC	480
QY	481	TGTGGCGTTTGGATCGCTTCGATGSGTTCGATGTTGAGTAATGCCGTTGATGCCCAATGA	540
DB	481	TGTGGCGTTTGGATCGCTTCGATGSGTTCGATGTTGAGTAATGCCGTTGATGCCCAATGA	540
QY	541	AGACCCAAAGCTTTACCTCCACCTGGCTTTTCACAGCTACATTTATTTGCTGGTGTTC	600
DB	541	AGACCCAAAGCTTTACCTCCACCTGGCTTTTCACAGCTACATTTATTTGCTGGTGTTC	600
QY	601	GGCTGCTTGGGTCTGTTAGGTTGGGTTGATCGTGGATTTTCTGTCAATGCCAT	660
DB	601	GGCTGCTTGGGTCTGTTAGGTTGGGTTGATCGTGGATTTTCTGTCAATGCCAT	660
QY	661	AATAGGGTTTCATGGGAGGAGCAGCCAGGTGGTGTCTGTCAGCAACTAAAATCGATTCT	720
DB	661	AATAGGGTTTCATGGGAGGAGCAGCCAGGTGGTGTCTGTCAGCAACTAAAATCGATTCT	720
QY	721	TGSCCTTGAGCAATTCACCCATGGAGCTGATATCATATCAGTGAAGCGCTCTGTTTTAC	780
DB	721	TGSCCTTGAGCAATTCACCCATGGAGCTGATATCATATCAGTGAAGCGCTCTGTTTTAC	780
QY	781	CCAACTCATGAGTGGAGTGGGAAAGTCTGTGTTAGGATGCTCTCAATTTCTTCTCT	840
DB	781	CCAACTCATGAGTGGAGTGGGAAAGTCTGTGTTAGGATGCTCTCAATTTCTTCTCT	840
QY	841	CCTTAGCACAAGATATCTCAGCAAAAAACGACCAGGTTTTTTTGGGTGTGACGAATGC	900
DB	841	CCTTAGCACAAGATATCTCAGCAAAAAACGACCAGGTTTTTTTGGGTGTGACGAATGC	900
QY	901	GCCATTGACGTCGCTTATATGGGAAGTCTCTTGGTTATTTTCACTCAGCCGAGAGCA	960
DB	901	GCCATTGACGTCGCTTATATGGGAAGTCTCTTGGTTATTTTCACTCAGCCGAGAGCA	960
QY	961	CGGTGTTGAAGTATAGGAAGTGAAGAGGTTTGAATCCACCATCTCAAAATCT	1020
DB	961	CGGTGTTGAAGTATAGGAAGTGAAGAGGTTTGAATCCACCATCTCAAAATCT	1020
QY	1021	GGTATTTGTGTCCTTTACATGACTACAGCTGTCAAAACCTGGCATTTGCTTGGCATCAT	1080
DB	1021	GGTATTTGTGTCCTTTACATGACTACAGCTGTCAAAACCTGGCATTTGCTTGGCATCAT	1080
QY	1081	ATCATTGCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTTACAATAT	1140
DB	1081	ATCATTGCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTTACAATAT	1140
QY	1141	TGATGSCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTCACCTC	1200
DB	1141	TGATGSCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTCACCTC	1200
QY	1201	TTGCTACCTCACAAAGGACCATTTTTCGGCTTCGGCTGTGAACATAAAGCTGGATGCAA	1260
DB	1201	TTGCTACCTCACAAAGGACCATTTTTCGGCTTCGGCTGTGAACATAAAGCTGGATGCAA	1260
QY	1261	GACAGAGCTTCCAAATTAATGATGCTTGCAGTAAATTTGACATTTGTTTCTCTGAC	1320
DB	1261	GACAGAGCTTCCAAATTAATGATGCTTGCAGTAAATTTGACATTTGTTTCTCTGAC	1320
QY	1321	ACCTTTGTTCCATTACACTCCCTGGTGGTGTATCAGCTATTATCGTATCTGCAATGCT	1380
DB	1321	ACCTTTGTTCCATTACACTCCCTGGTGGTGTATCAGCTATTATCGTATCTGCAATGCT	1380
QY	1381	TGGACTCATAGATTATGAAGAGCCATCCATCTATTTAAAGTTGACAAATTTGACTTTGT	1440
DB	1381	TGGACTCATAGATTATGAAGAGCCATCCATCTATTTAAAGTTGACAAATTTGACTTTGT	1440
QY	1441	GGTGTGCATGAGTGCAATACATTGGCGTGGTCTTTTGGCAGTGTGTGAATTTGGCTTAGTCA	1500

DB	1441	GGTGTGCATGAGTGCAATACATTGGCGTGGTCTTTTGGCAGTGTGAATTTGGCTTAGTCA	1500
QY	1501	AGCTATTGTAATATCTGTACTTTCGGGTACTTCTATTATTATTCGCAAGCCAGACATTCTG	1560
DB	1501	AGCTATTGTAATATCTGTACTTTCGGGTACTTCTATTATTATTCGCAAGCCAGACATTCTG	1560
QY	1561	TTTGGGCAACATTCGAAATTTCTGTGATATACGAAATGTTGAGCACTATCAAAATGCAAA	1620
DB	1561	TTTGGGCAACATTCGAAATTTCTGTGATATACGAAATGTTGAGCACTATCAAAATGCAAA	1620
QY	1621	ACATGTTCTCGAATGCTAAATTTCTAGAGATTGATGCCAAATTTACTTTGCCAAATGCCAG	1680
DB	1621	ACATGTTCTCGAATGCTAAATTTCTAGAGATTGATGCCAAATTTACTTTGCCAAATGCCAG	1680
QY	1681	CTATTTAAGAGAAAGGATCACAAGTGGATTGATGAAGAAAGAAAGAAATTTAAAGCTAC	1740
DB	1681	CTATTTAAGAGAAAGGATCACAAGTGGATTGATGAAGAAAGAAAGAAATTTAAAGCTAC	1740
QY	1741	AGGGGAGACTAGTTTGCAGTATGTTTAAATCATATGAGTGTCTTTGGAACATTTGATAC	1800
DB	1741	AGGGGAGACTAGTTTGCAGTATGTTTAAATCATATGAGTGTCTTTGGAACATTTGATAC	1800
QY	1801	AAGTGGAAATAGTATGCTTTGAAAGAGGTGAAGAAATTAACAGAGAGAGAGCTACAGCT	1860
DB	1801	AAGTGGAAATAGTATGCTTTGAAAGAGGTGAAGAAATTAACAGAGAGAGAGCTACAGCT	1860
QY	1861	TGTTTTGCTCAATCTGTAAAGTGAAGTGAAGAAATCTGAACAAATCGAAGTTCCTAAA	1920
DB	1861	TGTTTTGCTCAATCTGTAAAGTGAAGTGAAGAAATCTGAACAAATCGAAGTTCCTAAA	1920
QY	1921	TCATTTAGGAAAGAAATGGAATCTATCTGACTCTTCAAGAGGCGCTTTGGAGCATGCAACT	1980
DB	1921	TCATTTAGGAAAGAAATGGAATCTATCTGACTCTTCAAGAGGCGCTTTGGAGCATGCAACT	1980
QY	1981	CAATCTAGCTGCAAGCAAAACGAAACCAAGAAAGATGAACAGAGGGTTGAAACAATGT	2040
DB	1981	CAATCTAGCTGCAAGCAAAACGAAACCAAGAAAGATGAACAGAGGGTTGAAACAATGT	2040
QY	2041	GTGACTGATGCTATGCCAAGAGATTCTTAATTAATCTCAAAAGCTTATTCGTTTCTG	2100
DB	2041	GTGACTGATGCTATGCCAAGAGATTCTTAATTAATCTCAAAAGCTTATTCGTTTCTG	2100
QY	2101	CTTTAGTAATGTTTACCACTACAAATGTGTCATGAGAAATTTCTGAAATCAGCCGAAAGAT	2160
DB	2101	CTTTAGTAATGTTTACCACTACAAATGTGTCATGAGAAATTTCTGAAATCAGCCGAAAGAT	2160
QY	2161	TTTAAAGGCATAGGAAATGAAAGATGCAAGGGTCTTCTAAATTTCTCAACTCTGCATCT	2220
DB	2161	TTTAAAGGCATAGGAAATGAAAGATGCAAGGGTCTTCTAAATTTCTCAACTCTGCATCT	2220
QY	2221	TAGTTAGAGAAATCTCCTATGATAGGCTGTGAAATTAATCTTTACGTATCATGCTTG	2280
DB	2221	TAGTTAGAGAAATCTCCTATGATAGGCTGTGAAATTAATCTTTACGTATCATGCTTG	2280
QY	2281	ATAATATATTCAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACACATG	2340
DB	2281	ATAATATATTCAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACACATG	2340
QY	2341	TAAAGAGTAAAGAGTGTGTCAGCATCTCCTCCATATTCATTTGAAAGTAAATTTGCC	2400
DB	2341	TAAAGAGTAAAGAGTGTGTCAGCATCTCCTCCATATTCATTTGAAAGTAAATTTGCC	2400
QY	2401	ATGAGAAATTTAAATCTCTTTGGAAAAAATTTTGGAAAAAATTTTGGAAAAAATTTTGG	2449
DB	2401	ATGAGAAATTTAAATCTCTTTGGAAAAAATTTTGGAAAAAATTTTGGAAAAAATTTTGG	2449

RESULT 2  
US-09-720-317A-3  
; Sequence 3, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:

APPLICANT: Stephen M. Allen  
APPLICANT: Saverio C. Falco  
APPLICANT: Catherine J. Thorpe  
TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
FILE REFERENCE: BB-1167  
CURRENT FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/092,833  
PRIOR FILING DATE: 14-07-1998  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 3  
LENGTH: 1981  
TYPE: DNA  
ORGANISM: Zea mays  
US-09-720-317A-3

Query Match 31.2%; Score 765.2; DB 4; Length 1981;  
Best Local Similarity 65.8%; Pred. No. 4.3e-202;  
Matches 1112; Conservative 0; Mismatches 578; Indels 0; Gaps 0;

313 GAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCTCAGGGCATCAG 372  
7 GGAGTCCGACCTGATCGCGGCATCACATCGCCAGCTCGCCATCCCGCAGGGCATCAG 66  
373 TTATGCCAAGCTGCCAACTCCCTCCAATCTTGGACTATATNTTCAGACTTTATACACC 432  
67 CTAGCCCAAGCTGCCAACTCGCGCCGCTCGCTCGGACTCTACTCGAGCTTCGTGCGGCC 126  
433 ATTGATTATGCGATGATGGTAGCTCGAGGGATTTGGCAGTGGGACTGTGGCGTTGG 492  
127 GCTGGTGTAGCGCTGATGGGAGCTCAAGAGCTCGCGGTGGGAGCGGTGGCGTTGG 186  
493 ATCGCTTCTGATGGTTGATGTTGAGTAATGCGGTTGATCCCAATGAAGACCCAAAGCT 552  
187 GTCGCTCATCAGCTCCATGCTCGCAGCAGAGTGTGCCGACGAGAACCCCGTGTCT 246  
553 TTACTTCCACTGGCTTTCAGCTACATTAATTTGCTGGTGTGTTTTCAGGCTGCTTGGG 612  
247 CTACTTGCACCTCGCTTTCACCGCACCTTCTTGGCGGGGTCTTCCAGGCTCGCTCGG 306  
613 TCTGTTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 672  
307 CTTCTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 366  
673 GGGAGGAGCAGCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 732  
367 GCGCGGCGCGGAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 426  
733 TTTACCCATGGAGCTGATATCATATCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 792  
427 CTTTACCCATCCACCGACGCTGCTCGCTCATGGAAATCGGTCTTTCAGCCAGACACCA 486  
793 GTGAGTGGGAAAGTCTGTTAGGATGTTCTTCTCATTTCTTCTCTCTCTCTCTCTCTCT 852  
487 GTGCGGTGGGAGCGCTCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 546  
853 ATACTTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGCATTTGACGTC 912  
547 CTTTATCAGCAAGAGGCTCCCAAGCTGTTCTGGATCTCGCGGCGCGCGCTTGGAGTTC 606  
913 CGTTATTTGGGAGTCTCTTGGTTTATTTTACCTCAGCGCGAGAGCAAGCGGTGTTGAAGT 972  
607 CGTCTGCTCGGAGCGCTTCTGGTGTACCTCAGCGACGCTGAAACCCACCGGCATCGAAGT 666  
973 GATAGGAGAACTGAGAGGTTTGAATCCACCATCACTCACAAATCTGGTATTTGTGTC 1032  
667 GATCGGTTACTGAAAGAGGCTGAAATCCACCGTGGTGGAGAGCGCTGCAATTTCTACC 726  
1033 GCCTTACATGACTACAGCTGTCAAACTGGCATTTGCTGGGCATCATATCACTTTGCGGA 1092  
727 GCCCTACATGATGCTCGCGCTCAAGACTGGGATCATCACCGGGTCTATTCCTCTCGCGA 786

RESULT 3  
US-09-720-317A-15  
; Sequence 15, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe

1093 AGAATAGCAGTAGGAGAAAGCTTTGCAATGATATAAAAAATACATATTGATGCAACAA 1152  
787 AGAATAGCAGTAGGAGAAAGCTTTGCAATGATATAAAAAATACATATTGATGCAACAA 846  
1153 AGAGATGATAGCTATTGGGACCAATGAACGATGTTGGTCTTTTACCTCTTGTCTACCTCAC 1212  
847 GGAGATGATGCGGATCGGAGCGATGAACGCTCTGGGCTCGCTACGCTGCTGCTACCTGAC 906  
1213 AACAGGACCAATTTTCGGCTTCGGCTGTGAACATATAACGCTGGATGCAAGACAGCAGCTTC 1272  
907 CACGGGGCCCTTCTCGCGCTCCCGCGTGAACACTAACCGCGGGTGCAGGACGGCCATGTC 966  
1273 CAACATTAATGATGCTGCTGCACTTGCAGTAATGTTGACATTTGTTTCTGACACCTTGTTC 1332  
967 GAAAGTGGTCAATGCTGCTGCGGTGATGGTCAAGCTGCTGCTGCTGCGCGCTGTTTCCA 1026  
1333 TTACACTCCCTCGTGGTGTATCAGCTATTATCGTATCTGCAATGCTTGCAGCTCATAGA 1392  
1027 CTACACCGCGCTGGTGGTGTCTGCGGATCATGCTCTCGCGATGCTGCGGCTGCTGCGA 1086  
1393 TTATGAGCAGCCATCCATCTATTAAAGGTGACAAATTTGACTTTGTTGTTGTTGCTGATG 1452  
1087 CTTTGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146  
1453 TGCATACATTTGGGCTGCTTTCGCGAGTGTGAAATTTGGCTTAGTATGCTATGCTATTTAAT 1512  
1147 CGGTAACCTGGGCTGCTTTCGCGAGCTGCGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTG 1206  
1513 ATCTGACTTTCGGGTACTTCTATTATTTGCAAGGCCAAGGACATTTGTTTTGGGCAACAT 1572  
1207 CTCCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266  
1573 TCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACATGTTCTCTG 1632  
1267 CCCCGCACCATGCTGTACCGAGGATGGAACAGTACGCGCGCGCGCAGACGCTGCTGCTG 1326  
1633 AATGCTAATTTCTAGATTTGATGCAACAAATTTTACTTTGCAATGCGAGCTATTTAAGAGA 1692  
1327 CGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1386  
1693 AAGGATCACAAAGTGGATTTGATGAAGAAAGAAAGAAATTTAAAGCTACAGGGAGACTAG 1752  
1387 GAGGATCTCGCGTGGATGCAACGAGGAGGAGCGCAACAGAGCCAGGCGAGATGGG 1446  
1753 TTTGCAAGTATTTAATTTGATATGATGCTGTTGGAAACATTTGATACAAGTGGAAATAG 1812  
1447 CGTGGGTAGCTTGTCTCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506  
1813 TATGCTTGAAGAGTGAAGAAATTTACAGAGAGAGAGCTACAGCTTGTGTTGTTGCTCAA 1872  
1507 CATGCTGGAGAGCTCAACAAAGTCTTTGGAACAGGGGAAATGCAATGCTGCTGCTGCTG 1566  
1873 TCCGTGAAGTGAAGTGAAGAAACCTGAACAAATCGAAGTTCCTCAAAATCATTTAGGGAA 1932  
1567 CCGGCGAGGAGATCATGAAGAGCTGGAGAGCTCAAGGCTGCTGAGGAGATTCGGCCA 1626  
1933 GAAATGGATCTATCTGACTGTTGAAAGAGCGCTTGGAGCATGCAACTTCAATCTACGTGC 1992  
1627 CGAGTGGGTGTTCCGACAGTGGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1686  
1993 AGCAAAACG 2002  
1687 GCACAAAGCG 1696

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins									
; FILE REFERENCE: BB-1167									
; CURRENT APPLICATION NUMBER: US/09/720,317A									
; CURRENT FILING DATE: 2000-12-21									
; PRIOR APPLICATION NUMBER: 60/092,833									
; PRIOR FILING DATE: 14-07-1998									
; NUMBER OF SEQ ID NOS: 31									
; SOFTWARE: Microsoft Office 97									
; SEQ ID NO 15									
; LENGTH: 2067									
; TYPE: DNA									
; ORGANISM: Glycine max									
US-09-720-317A-15									
Query Match 29.4%; Score 720.6; DB 4; Length 2067;									
Best Local Similarity 62.2%; Pred. No. 1.1e-189;									
Matches 1134; Conservative 0; Mismatches 689; Indels 0; Gaps 0;									
QY	143	CGCCACAGCCGTTTTCAAGTCTCTAAAGTACTCTTTGAAGGAGACTTTTCTTCCCTGAT	202						
DB	8	CCACACCAGACACACTCCACAAACTTAGCACAGAGTCTCCGAATCTTCTCCAGAT	67						
QY	203	GACCTTTGAGGAGTTCAAGAAACAGCAGCTTCCAAGAGTTCAATGCTTGGCCTTCAG	262						
DB	68	GACCTCTCCACCGTTTTCAAGAACCAAACTCGCTTTAAAAAGTTCTCTCGCACTTCAG	127						
QY	263	TTCTTCTCCCAATTTTCCGAATGGGCTCCCAATACACCTTTCAGTTCTTGAAGCTGAC	322						
DB	128	TATCTCTTCCCAATTTTCCAGTGGGCCCAAACTCAATCTTACCCTTCTCCGCTCTGAC	187						
QY	323	CTCATAGCTGGCATCACTCGTAGCTTGGCCATTCCTCAGGGCATCAGTTATGCCAAG	382						
DB	188	CTCATCTCTGGCCTCACCATTGCCAGCTCCGCAATTCCTCAGGGAATCAGTTATGCCAAG	247						
QY	383	CTCGCCAACTCCCTCCAAATCTTGACATATATTCGAGCTTTATACCAATGATTTAT	442						
DB	248	CTTGCCAACTTGCACCTATTTCTGGAATATATTCGAGTTTGTTCCTCCCAATGATATAC	307						
QY	443	GCATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGACTGTGGCGTTTCGATCGCTTCG	502						
DB	308	TGCTGCTTGGAAAGTTCTAGACATCTTGGTTGGACCTTGTTCATTCGGTCTTTGGTTC	367						
QY	503	ATGGGTTGATGTGAGTAATCCGCTTGATCCCAATGAAGACCCAAAGCTTTACCTCCAC	562						
DB	368	ATGGGATCAATGTTAAGTGATATAAATTTCTTACACTCAAGAACCTATTCTCTATCTGGGA	427						
QY	563	CTGGCTTTACAGCTACATATTTGCTGGTGTGTTTTTCAGGCTGCCCTGGGCTGTTTAGG	622						
DB	428	TTGGCTTTTACCGCCACTTTCTTTGCTGGTGTATTCGAAGCTTCTCTGGGTATATTTAAGG	487						
QY	623	TTGGGTTGATCGTGGATTTTCTGTACATGCCAACCATAATAGGGTTTCATGGGAGGCA	682						
DB	488	CTAGGCTTCGTAATGATTTTCTGTGAAGGCACGCTGGTGGATTCACGCGGTGCT	547						
QY	683	GCCACGGTGGTGTCTGCAGCAACTATAATCGATTTCTTGGCCTTGAGCAATTTCACCCAT	742						
DB	548	GCCATTTATTTGTCTCTGCAGCAGCTGAAGGTTTACTTTGGAATAGTGCACTTTTACCGG	607						
QY	743	GGAGCTGATATCATATCAGTATGCGCTGTGTTTTTCAACCAACTCATGAGTGGAGTGG	802						
DB	608	AAGATGCAATAATTTCCGATTAACGATCTCTGTGTTTTTCAAGCAAGACACAGTGGTCATGG	667						
QY	803	GAAAGTGTGTGTAGGATGTGCTTCATTTTCTCTCTAGCAACAAGTACTTTCAGC	862						
DB	668	CAAAACATTTTGGGATTCGGCTTCTGGTCTTCTGCTGCAACAAGGCACATTAGT	727						
QY	863	AAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGGCCCATATGACGTCGGTTATATTG	922						
DB	728	TTGAGGAACCAAAACTATTCTGGGTTTCAGCAGCTGCCCATTTGACATCAGTTATTCTG	787						
QY	923	GGNAGTCTCTGGTTTTATTTCACTACGCGGAGACGAGGTTTGAAGTGATAGAGAA	982						
DB	788	TCAACCAATTTTAGTCTTTCTTCTGAGAAATAAGACTCATCAAAATTTTCAGTTATTGGGCAC	847						

; Patent No. 6696292  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stephen M. Allen  
 ; APPLICANT: Saverio C. Falco  
 ; APPLICANT: Catherine J. Thorpe  
 ; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
 ; FILE REFERENCE: BB-1167  
 ; CURRENT APPLICATION NUMBER: US/09/720,317A  
 ; CURRENT FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/092,833  
 ; PRIOR FILING DATE: 14-07-1998  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 1  
 ; LENGTH: 2279  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 US-09-720-317A-1

Query Match 26.8%; Score 655.4; DB 4; Length 2279;  
 Best Local Similarity 60.4%; Pred. No. 1.6e-171;  
 Matches 1120; Conservative 0; Mismatches 726; Indels 9; Gaps 2;

QY	121	AGTGACCAAGTCGAGGTTCCACCGCCACACGCGTTTTCAGTCTCTAAAGTACTCTTT	180
DB	188	AGTGCACAAGGTGGCGCGCCACCGCGCGGAGCAGCGGAGATGAAGGTGAGGGT	247
QY	181	GAAGGAGACTTTCTTCCCTGATGACCCCTTTCAGGCAGGTTCAAGAACAAAGCCAGCTTCCAA	240
DB	248	GAAGGAGACTTCTTTCCTCCGAGACCCGTTCCGGGCGGTTCAAGGGGCGAGCCGCGGGGAC	307
QY	241	GAAGTTTCATGCTTGGCCCTTCAGTCTCTTCTCCCAATTTTCCAAATGGGCTCCCAATACAC	300
DB	308	GCAGTGGCTCATGGCGGTTCAGGTACCTCTTCCCCATCTCGACTGGGTGCGGACTATCTC	367
QY	301	CTTTTCAGTTCTTGAAGCTGACCTCATAGCTGGCATACCATCGTAGCTTGGGCATATCC	360
DB	368	CTTGTGCTCTCTTCAAGTCCGACCTCGTCGGGGGCTCCACCATTGCAGCCTCGCCATATCC	427
QY	361	TCGGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCCAATTTCTTGGACTATATTCGAG	420
DB	428	TCAGGGCATTTAGCTACGGGAAGCTTGGCAAGCTGCTCTCCCAATATCGGGCTGTATTCGAG	487
QY	421	CTTTATACCAACCATTTGATTTATGCGATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGAC	480
DB	488	CTTCGTCCCGCGATGGTGTACGCGGTGCTGGGAGCTCCCGCTGACCTGGCGGTGGGGCC	547
QY	481	TGTGCGGTTGGATCGCTTCTGATGGGTTTGATGGGTGATGTTGAGTAAATGCCGTTGATCCCAATGA	540
DB	548	GGTGTGATCTCTGTCGTGATCATGCGGTTCATGCGGTTCATGCTGCGGAGCGCTGAGCCCCACTGC	607
QY	541	AGACCCAAAGCTTTTACCTCCACCTGGCTTTCCACAGCTACATTTATTTGCTGCTGTTTTTCA	600
DB	608	GGAGCCGACGCTGTCTCTGCACTGGCCCTTCACTTCACCTGTTGCGGGGCTGTGTGCA	667
QY	601	GGCTGCTCTGGGCTGTGTTAGTTTGGGTTGATCGTGAATTTCTGTCAATGCAATCCAT	660
DB	668	GGCTCTCTGGGCATCCTCAGGCTCGCTTGGTTCATCGACTTCCTGTCCAAGGCGACGCT	727
QY	661	AATAGGTTTCATGGGAGAGCAGCCACGGTGGTGTGTCTGAGCAACTTAAATCGATTCCT	720
DB	728	GGTGGGGTTTCATGCCCGCGCGCCATCATCGTGGCGCTGCAGCAACTCAAGGGGCTGCT	787
QY	721	TGGCTTTGACATTTCAACCATGGAGCTGATATCATATCATGATGCGCTCTGTTTTTCAC	780
DB	788	GGGCATCGTCCACTTCACCAACCGGAGATGGGCAATCGTCCAGTTCATGGCCTCCGCTTCCCA	847
QY	781	CCAAACTCATGAGTGGAGGTGGGAAAGTGCTGTTAGGATGTGTCTTCAATTTCTTCTCT	840
DB	848	CCACACGAGGATGGTGTGTGGCAGACGATCCTCATGGGCGCTGCTTCTCTGCTTCTCT	907
QY	841	CCTTACGACAAGATATTTACGAAAAAACGACCAAGGTTTTTTTTGGGTGTACGAAATGGC	900

908	908	908	GCTGTGGCGAGGATGTGAGCATCAGATGGCCAAAGCTTTTCTGGGTTTCGGCGTGGCGG	967				
Qy	Db	Qy	GCAATTCAGCTCCGTTTATATATGGGAAGCTCTCTTGGTTTATTTCACTCACGCCGAGAACGA	960				
					Db	Qy	GCCCCGGCATCGGTCAACATCTCGACGCTGCTGTGTTTCTCTTCAAAGCTCAGAACCA	1027
Db	Qy	TGGCATCAGCATCATTTGGCGAGCTCAAGTCGGCGCTGAATCGCCCTCTGTTGGACAAAGCT	1087					
				Db	Qy	GGTATTGTCTCGCTTTACATGACTACAGCTGTCAAAACCTGGCATTTGTCTGGCATCAT	1080	
Db	Qy	CCTGTTTGAACGCGCGTATTTAGGCCTCACCATGAAGACTGGCCCTGTGCACCGGAATCAT	1147					
				Db	Qy	ATCACTTCCGGAAGGAATAGCAGTAGGAAGAAGCTTTTGCAATGTATATAAAATTTACAATAT	1140	
Db	Qy	CTCACTGACGGAAGGAATAGCGGTTGGTAGAACATTTTGCCTCACTCAAGGACTACCAGAT	1207					
				Db	Qy	TGATGGCAACAAAGAGATGATAGCTATTTGGGACATGAAGCAAGTATAGTTGTCTTTCACCTC	1200	
Db	Qy	AGATGGAAAACAGGAGATGATGGCCATAGGCTTGAATGTTTGTGGTCTCTGCACATC	1267					
				Db	Qy	TTGCTACTCACAACAGGACCAATTTTCGGCTTCGGCTGTGAATCATTAACGCTGGATGCAA	1260	
Db	Qy	ATGCTAGCTAAACACAGTGGCTTCTCCGCTCTGCTGTAAACCAACGCCGGCTGCAA	1327					
				Db	Qy	GACAGCAGCTTCCAAACATTAATATGTCATCTGCAGCTAACTGTGACATTTGTTATTCCTGAC	1320	
Db	Qy	GATGCCATGTCNAACGTGATCATGGCGCTGACTGTGATGGTCAAGCTGCTGTCTCTCAT	1387					
				Db	Qy	ACCTTGTGTCATTAACACTCCCCTGGTGGTGCTCATGCTATTTATCGTATCTGCAATGCT	1380	
Db	Qy	GCCACTGTTGCTGTACACACCCAAAGTTGTCTCGGAGCGATCATCATGCCCGGTGAT	1447					
				Db	Qy	TGGACTCATGATTAAGAAGCAGCAATCCATCTATTTTAAAGTTGACAAATTTGACTTTGT	1440	
Db	Qy	CGCCCTGATGCAATTTCCCGCGGTGACCACTCTGGAAGATGGAACAAGATGGATTTCT	1507					
				Db	Qy	GGTGTGATGAGTGTCATACATTCGGCTGGTCTTTGGCAGTGTGTAATTTGGCTTTAGTCAT	1500	
Db	Qy	GGTGTGGGTTTGGCGTTTTCGGCGCTCATCTTCTCATCTCAGTCCAGAGGCCCTTGGAT	1567					
				Db	Qy	AGCTATTTGTAATATCTGTACTTCGGGTACTTCTTATTTATTTGAAGGCCAAGGACATTCGT	1560	
Db	Qy	AGCGTTGGTATATCTATATTTAGGCTGTGATGTCAGATCACAAGGCCGAAGATGATGGT	1627					
				Db	Qy	TTTTGGCAACATTTCCAAATCTGTGTATATCCGAATTTGTGACACTATCAAAATGCAAA	1620	
Db	Qy	TCAAGGGAACATCAAGGGGACTGATATTTACAGAGACCTGTCATCTACTCAAGGAGGCCCA	1687					
				Db	Qy	ACATGTTCTCGGAATGCTAATCTCAGAGATTGATGACCAAAATTTACTTTTGCCAAATGCGAG	1680	
Db	Qy	AAGAGTTTCTGGGTTCTTGATCTTTGGCCATTGAAGCACCGATAAATCTTGCCAACTCCAA	1747					
				Db	Qy	CTATTTAAGAGAAAGGATCAACAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTAC	1740	
Db	Qy	CTACCTGAATGAAGGATTAAGAGATGGAT-----AGAGGAAGAAATCTTTTGAACAGGA	1801					
				Db	Qy	AGGGGAGACTAGTTTGCAGTATGTTATTAATGATGAGTCTGTTGGAAACATTGATATC	1800	
Db	Qy	TAAACATCTAGAACTCCATTTTCATTAATCTTTGGATCTGTACAGCTGTGTTCTGCAATTTGACAC	1861					
				Db	Qy	AAGTGAATTAAGTATGCTTGAAGAGGTGAAGAGATTTACAGAGAGAAGAGAGCTACAGCT	1860	
Db	Qy	AAGTGGCATAGCGTTCTCATTTGACATTAAGAAATCAATAGAGAAACGTGGTCTGGAGCT	1921					
				Db	Qy	TGTTTTGGTCAATCTGTAAGTGAAGTGAAGAAATCTGAAACAAATTCGA---AGTTTCCA	1917	
Db	Qy	TGTTGCTGTCAATCCAATCGGAAGCTCATGGGAGAAATACACAGTGCAAACGAGGCTGA	1981					
				Db	Qy	AAATCAATTTAGGGAAGAAATGGAATCTATCTGACTGTGTGAAGAGCGGCTTGGAGCA	1972	
Db	Qy	AAACTATTTTGGCCAGATTCGTTGTATCTGACACTGGCGAAGCAATCGCTTCA	2036					

RESULT 5

US-09-720-317A-19  
; Sequence 19, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saviero C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR FILING DATE: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 19  
; LENGTH: 2311  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-09-720-317A-19

Query Match 26.1%; Score 640.4; DB 4; Length 2311;  
Best Local Similarity 59.4%; Pred. No. 2.3e-167;  
Matches 1104; Conservative 0; Mismatches 751; Indels 3; Gaps 1;

RESULT 5																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



Db 1927 CTCATTCTGTCGAATCTCGTTCGGTCCGTATAGAAAACTGCAAGCGTCGAAGCTCACC 1986  
Qy 1919 AATCAATTAGGAAGAATAATGATCTATCTGACTTTGAAAGGCGGTGGAGCATGCA 1976  
Db 1987 GAGCACATTGGAAGCAGCAATATATCTCGCGGTCTCTGACGCTGTGCGATTCTGTA 2044

RESULT 6  
US-09-720-317A-7  
; Sequence 7, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 780  
; TYPE: DNA  
; ORGANISM: Helianthus tuberosus  
US-09-720-317A-7

Query Match 13.2%; Score 324; DB 4; Length 780;  
Best Local Similarity 70.6%; Pred. No. 1.2e-79;  
Matches 432; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

Qy 1400 GCAGCCATCCATCTATTAAAGTTGACAAATTTGACTTTGTGTGTCATCAGTGCATAC 1459  
Db 17 GCAGCGATTACCTCTCGACACATAGACAAATTCGACTTTGTGTGTCATGAGTGCATAC 76  
Qy 1460 ATTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCTATAGTATTTGTAATCTGTA 1519  
Db 77 TTTGGTGTGTCTTTGGAGTGTGAAATTTGGATTAGTATTCGCGGTGCGATTGTCTGTG 136  
Qy 1520 CTTGGGTACTCTCTATTATTGCAAGGCCAGGACATTCGTTTGGCACAATTTCCAAAT 1579  
Db 137 CTTAGGGTACTCTCTATTATTGCTCGAGGCCAAGAACATCGACGCTAGTCTCATACCGAT 196  
Qy 1580 TCTGTGATATACCGAAATTTGAGCACTATCAAAATGCAAAACATGTTCTCGAATGCTA 1639  
Db 197 TCCACTATTATAGAAGTATGGATCAATACCAAAATCGGAAGCGTTCCAGNACTTG 256  
Qy 1640 ATTCTAGAGATTGATGCCAAATTTACTTTGCCAATGCCAGCTATTTAAGAGAAAGATC 1699  
Db 257 ATACTTCAAATCGAAGCACCTATTATTCTTTGCTAACTTAGCTACTTTGAGGAAAGGATT 316  
Qy 1700 ACAAGTGGATTGATGAAGAGAGAAAGAAATTAAGCTACAGGGAGACTAGTTTGCAG 1759  
Db 317 GTGAGATGGGTGATGAAGAGAGAGATAGGTGGAAGTCTTTAAGGAGAAATGACTTGCAA 376  
Qy 1760 TATGTTAATAATGATATGAGTGTCTTTGGAAACATTTGATACAAGTGAATTAAGTATGCTT 1819  
Db 377 TATGTCAATCTTGCAATTGAGTGTCTTTGGAAATATTGATACAAGTGGNATAACATGCTT 436  
Qy 1820 GAAGAGTGAAGAAGATTACAGAGAGAGAGAGCTACAGCTTGTGTTTGGTCAATCTCTGTA 1879  
Db 437 GGAGAAGTTAAAGAGTTATGGAAGAGAGAGGCTAAAGTTGGTTTGTAGCGAATCCGGGC 496  
Qy 1880 AGTGAAGTGAAGAAGTGAACAAATCGAAGTTCCAAATCATTTTAGGGAGAAATGG 1939  
Db 497 GGAGAGTGAAGAAGATGAACAAAGCGAAGTTGATAGAGGTGATCGGGCAGAAATGG 556  
Qy 1940 ATCTATCTGACTGTGTAAGAGGCGGTGGAGCATGCAACTTTCAATCTACGTGCAAGCAAA 1999  
Db 557 ATATATCTACAGTGGGAGAGCGGTGGAGCGTGCACCTTTATGCTTTCATCTTACAG 616

Qy 2000 ACGAACCCAAAG 2011  
Db 617 AACGCCGAAAAG 628

RESULT 7  
US-09-720-317A-5  
; Sequence 5, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 1240  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-720-317A-5

Query Match 8.8%; Score 216.2; DB 4; Length 1240;  
Best Local Similarity 53.4%; Pred. No. 1.4e-49;  
Matches 477; Conservative 0; Mismatches 413; Indels 3; Gaps 1;

Qy 1085 CTTGCGGAAGGAATAGCAGTAGGAAGAAAGCTTTGCAATGTATAAAAATTAACAATTTGAT 1144  
Db 9 CTCACGGAAGCTATCGCGTTGGCCGATCTTTCGCCCTCGGTAAAGGGGTACAGACTCGAC 68  
Qy 1145 GGCACAAAGAGATGATAGCTATTGGGACCATGAAAGTAGTTGGTTCTTTTCACTCTTGC 1204  
Db 69 GGCACAAAGAGAGATGTGGCCATGCGGGTCTCCAAAGTGTGCTGGTTCTCTGTCTCTGTC 128  
Qy 1205 TACTCTCAACACAGGACCATTTTCGCTTGGCTGTGAACCTATATAACGTGTGATCAAGACA 1264  
Db 129 TATGTGGCAACAGGTTCGTTTCGCCGAA CGGCAAGTAACTTCAGCGGGGGCCAGGTGC 188  
Qy 1265 GCAGCTTCCAAACATTAATATGTCATCTTGCAAGTAATGTTGACATGTTATTCCTGACACC 1324  
Db 189 ACCGTTTCAACATCGTCTCATGTCATCACCGTTCCTCACCTGGAGCTGTTTCATGAAG 248  
Qy 1325 TTGTTCCATTACACTCCCTGGTGGTGTCTATCAGCTATTATCGTATCTGCAATGCTTGA 1384  
Db 249 CTCCTCTACTACAGCCCATGCGGTGCTCGCCCTCCATCATCTCTCGGCTCTTCGGGA 308  
Qy 1385 CTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTGTGGTG 1444  
Db 309 CTGATCGACATCAAGGAGCGCTGCAGCATATGGAAGATCGACAAGATGATTTCTCACC 368  
Qy 1445 TGCATGAGTGCATACATTCGCGTGGTCTTTGGCAGTGTGTAATTCGCTTAGTCATAGCT 1504  
Db 369 TGCCTCGGTGCGTTCGCGTTCCTGTTGGTTCGCGTGGAGATTCGGCTTGCAGTTGCA 428  
Qy 1505 ATTGTAATATCTGACTCTCGGTACTTCTATTATTATTCGAAGGCCAAGGACATTCGTTTG 1564  
Db 429 CTTGGCATTTCTTCGCAAGATCATCATCAGTCTCGGCTTCAGGTGGAGATCTCTT 488  
Qy 1565 GGCACATTTCCAAATTTCTGTGATATACCGAAATTTGAGCACTATCAAAATGCAAAACAT 1624  
Db 489 GGCAGCTACAGGAGACATATCTTCTGACGCTCAGGACGTACCTGTAGCTGCCTA 548  
Qy 1625 GTTCTCGAATGCTAATCTTAGAGATGTGATGACCAATTT---TACTTTGCCAATGTCAGC 1681  
Db 549 ACTCCGACTGACTGCTATAGCGTCGACACATCTCTCTCTGCTTTCATCAACGCCACT 608  
Qy 1682 TATTTAAGAGAAAGGATCAAGGTGGATTGATGAAGAGAAAGAAATTAAGCTACA 1741



Db 609 TCCGTCACAAAGAAAGGATCACAGAGTGGGTTTGGGAAGGAGTGGAGACCTCAAAATGGAAAA 668  
Qy 1742 GGGGAGACTAGTTTGGCAGTATGTTAATAATTGATGATAGTCTGTTGGAACATTTGATACA 1801  
Db 669 GCAGGGAGAGGATACAGCAGTTGCTTGATATGTCAGTGTGTTAAACATCGACACT 728  
Qy 1802 AGTGGATAAGTAGTCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTT 1861  
Db 729 TCAGGACTCACTGCACTGGGAAGAAATACACAAGGAGTTGGTGCTCTTGGCTTACAGATG 788  
Qy 1862 GTTTTGGTCAATCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCAAAT 1921  
Db 789 GCTATAGCCAGTCCGGATGGAAGGCAGTTTCAAGAAGATGAAGTGTACAGAGTGGTGGAC 848  
Qy 1922 CATTTAGGGAAGAAATCGATCTATCTGACTGTTTGAAGAGGCCGTTGAGCATG 1974  
Db 849 AGGTAGGACAGGACTGGATCTTCAATGACAGTAGGTGAAGCGGTGGAGGCGCTG 901

RESULT 8  
US-09-720-317A-21  
; Sequence 21, Application US/09720317A  
; Patent No. 6698292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720, 317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 21  
; LENGTH: 2022  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-09-720-317A-21

Query Match 8.4%; Score 205.6; DB 4; Length 2022;  
Best Local Similarity 47.0%; Pred. No. 1.7e-46;  
Matches 791; Conservative 0; Mismatches 869; Indels 24; Gaps 4;

Qy 306 AGTCTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTTCCCTCAGS 365  
Db 103 AGGACTTCCAGCGCGACCTCGCCGCGCATCACTGTGCGCGTCATGCTTGCCTCAGG 162  
Qy 366 GCATCAGTTATGCAAGCTCGCCAACTCCCTCCAATTTCTTGCACTATATTCGAGCTTTA 425  
Db 163 CAATGTCATATGCAAGCTGGCTGGGCTTCACCCAAATTTATGGCTCTACACAGGCTTG 222  
Qy 426 TACCACATGATTATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAGCTGTGG 485  
Db 223 TCCCACTATTTGTCTACGCGAATTTTGGGTCTCTCAGCAAAATAGCAGTAGGTCCAGTGG 282  
Qy 486 CGGTTCGATCGCTTCTGATGGGTTTCGATGTTGAGTAATCGGTTGATCCCAATGAGACC 545  
Db 283 CACTGTCTCTGTAGTGTCCAAATGTTCTGGGGGTATAGTTAATTCATCTAGTG--- 339  
Qy 546 CAAAGCTTTTACCTCCACCTGGCTTTTCACAGCTACATTAATTTGCTGTGTTTTTCAGGCTG 605  
Db 340 ---AGCTGTACACGGAATTAGCCATATATTGGCAATTCATGGTTGGAATCTGGAATGCT 396  
Qy 606 CTTGGGTCTGTTTAGTGGGTGATCGTGGATTTTCTGTCACATGCAACCATTAATAG 665  
Db 397 TGATGGCAITGTCTAAGACTTGGCTGCTTATTCGTTTCAITAGCCATTTCTGTAATATCTG 456  
Qy 666 GGTTCATGGGAGGAGCAGCACCGTGGTGTCTCGCAGCAACTAAATTCGATTCCTGGCC 725  
Db 457 GATTCTACAGCTTCGGCCATCTGTAATTTGGTTTGTGCCAAATCAAGTATTTCTTGGG--- 514

Qy 726 TTGAGCAATTTACCCATGAGCTGATATCATATCAGTGTGGCTCTGTTTTCCACCAA 785  
Db 515 -TTACAGTGTACAGAAGTAGCAAAATTTATACCCTTATTGAGAGTATAATTGCTGGAA 573  
Qy 786 CTCATGAGTGGAGGTGGGAAAGTGTGTGTAGGATGTCTTCAATTTCTTCTCCTCTTA 845  
Db 574 TAGATCAGTTCTCTGGCCTCCATTTGTAAATGGGATCAGCGTTTCTTGTATTCTTCTAA 633  
Qy 846 GCACAGATACCTTCAGCAAAACACGACCAAGGTTTTTTTTGGGTGTGACCAATGGCGCCAT 905  
Db 634 TAAAGAAAGCTAGGGAACAAATAAATAAATAAGTTTCTCTGAGAGCTTCTGTGTCAC 693  
Qy 906 TGACGTCCGTATATTGGGAAGTCTCTTGGTTTATTTTCACTCACGCCGGAAGACAGCTG 965  
Db 694 TAAACAGCTGTCTTCTTGGAACTTGTGTGAAATTTTCCGTCC-----AAGTCCCA 747  
Qy 966 TTGAAGTGTAGAGAGAACTGAAGAAGGTTTGAATCCACCATCACTCAAAATCTGGTAT 1025  
Db 748 TATCAGTGTAGGTGAAATACCCCAAGGCTTCCCAAGTTTCCCAATTTCTCCATTTCTCGAGGATTTG 807  
Qy 1026 TTGTGTGCGCTTACATGACTACAGCTGTCAAACTGGCAATTCGTGTGGCATCATATCAC 1085  
Db 808 AACATCTGATGCCCTAATGCCAACTGCAATACTTATCACTGCTGTG-----CTA 858  
Qy 1086 TTGCGGAAGGAATAGCAGTAGGAAGGCTTTTGCATATGTATAAAATTTACAATATTGATG 1145  
Db 859 TTTTGGAGTCTCTTGGGATTGCTAAAGCGTTAGCTGCGAAGAATGTTTATGATTGGACT 918  
Qy 1146 GCACAAAGAGATGATAGCTATTGGGACCAATGAAGTAGTGTGTTCTTCACTCTTGTCT 1205  
Db 919 CAAACAAAGAGTATTATTGGCCTTGGCTTATCAAAATATATGCGGTTTCAATTTCTCTCGCAT 978  
Qy 1206 ACCTCACACAGCAGCATTTTCGCTTGGCTGTGAACCTATAACGCTGATGCAAGACAG 1265  
Db 979 ATCTCTGTACAGCTCTCTTTCTAGGTCTGTGTGAATCATGAAAGCGGGGCAAGACTG 1038  
Qy 1266 CAGCTTCCAACATTAATATGTCATTTGCAAGTAAATGTTGACATTTGTTATTCCTGACACCT 1325  
Db 1039 GATTATCAGGAATCATATGGGCATAAATTTGCAGTGTCTCTTGTATTATGACACCAT 1098  
Qy 1326 TGTTCATTTACACTCCCTGTTGGTGTCTATCAGCTATTATCGTATCTGCAATGCTTGGAC 1385  
Db 1099 TATTTACTGATATACCTCAGTGTGCAATTTGGCTGCAATTTGTGATTTCTGCTCAGCTGGCC 1158  
Qy 1386 TCATAGATTATCAAGCAGCATCCATCTATTAAAGTTGACAAATTTGACTTTTGGTGTCT 1445  
Db 1159 TGGTAGATTGGAAGAGGCCATCTTCTGTGGGTATTGATGAAGAGGATTTCTTCTCT 1218  
Qy 1446 GCATGAGTGCAATTCAGTGGCTGCTTTTGGCAGTGTGAAATTTGGCTTTAGTCATAGCTA 1505  
Db 1219 GGGCGATGACATTTACTACAACCTTAACTTTTGGCAATTGAGATTGGTGTCTTGTGGGG 1278  
Qy 1506 TTGTAATATCTGTACTTCGGGTACTTCTATTATTATGCAAGGCCAAGGACATTCGTTTTGG 1565  
Db 1279 TCGGTTTTTCGCTGGCATTTTGTGATCCCATGAATTCGCAATCCGCATATAGCTGTTTTGG 1338  
Qy 1566 GCACCAATTCCTGATATACCGAAATGTTTGAGCACTATCAAAATGCAAAACATG 1625  
Db 1339 GCGTTTTGCTGGCACCACTGTGACAGGAATACATTCGAGTACCTTGAGGCTTATACAT 1398  
Qy 1626 TTCTTGGAAATGTAATTCAGAGATTGATGCAACCAATTTTACCTTCCCAATGCCAGCTATT 1685  
Db 1399 ACAACGGGATTTGTTGTGTCGTTGTGATGACCACTACTTCTTGTAACTAAGTTTACA 1458  
Qy 1686 TAAAGAAAGGATCACAAGGTGGATTGATGAAGAGAAGAAAGTAATAAGCTACAGGGG 1745  
Db 1459 TAAAGGACAGGTTGCTGAGTATGAGCTCAAACTCCCAATTCAAACCGTGACCTGATG 1518  
Qy 1746 AGACTAGTTTTCAGTATGTTAATAATTGATATCAGTGTCTTGGAAACATTTGATACAAGTG 1805  
Db 1519 TTGGAGGGGTGACTTTTGTGATCTCGAGATGTCCTCTGTTACATACATCGACTCGAGCG 1578  
Qy 1806 GAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTTGT 1865

Db 1579 CTGTTCAAGCTCTCAAGGAGCTGCAACCAAGATACAAAGCAGCGGACATCCAGATTGCTA 1638  
Qy 1866 TGGTCAATCTCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCAAATCATTT 1925  
Db 1639 TAGCGAATCTTAACCGGAGGTGACCTATTGCTGTCAAGAGCGGCATCATCGACATGA 1698  
Qy 1926 TAGGGAAGAATGATCTATCTGACTGTTGAGAGGCGGTGGAGCATGCAACTTCAATC 1985  
Db 1699 TTGGCGGAGGTGTTGTTCTCGTAGTGCACGACGCGGTGCAAGTATGCTTCAGCATG 1758  
Qy 1986 TACG 1989  
Db 1759 TGGC 1762

RESULT 9  
US-09-720-317A-13  
; Sequence 13, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR FILING DATE: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 13  
; LENGTH: 493  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-09-720-317A-13

Query Match 4.8%; Score 118; DB 4; Length 493;  
Best Local Similarity 55.0%; Pred. No. 1.7e-22;  
Matches 252; Conservative 0; Mismatches 205; Indels 1; Gaps 1;

Qy 816 TAGGATGTCTTCATTTTCTTCTCTTAGCACAGATCTTCCAGCAAAAGGACCA 875  
Db 36 TCGGATGCTCTTCTCTCATATTCATCTCCACACAGGTTTCATCGGAGAGGTACAA 95  
Qy 876 GGTTTTTTGGGTGTCAGCAATGGCGCATTTGACGTCCGTTATATTGGGAAGTCTCTGG 935  
Db 96 AGCTGTTCTGGGTGTCAGCATCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 155  
Qy 936 TTTATTTTCACTACGCGGAGAGCAGCGGTGTTGAAGTGTAGGAGAACTGAAGAAGGGTT 995  
Db 156 TCTACGGGACAAGGGGTGACAGGACGCGGTCAAGATCATCCAGAAGGTGCACGCGGCC 215  
Qy 996 TGAATCCACCATCTACT-CACAAATCTGTTATTTGTGTGCTGCTTACATGACTACAGTGT 1054  
Db 216 TAAACCCCAAGTCCGTGGAGAGATACACTCAACGCGGCGCACACAAACGAGTGGGCC 275  
Qy 1055 AAAATCGCATTTGCTGTCATCATATCTCTGCGGAAGGAATAGCAGTAGGAAGAC 1114  
Db 276 CAAGATCGCGTCTATCTGCGCATCATCGCCCTCAGGAAAGTATCGCCGTGGCCGATCT 335  
Qy 1115 TTTTGAATGTATAAAATTAACAATTTGATGGCAACAAAGAGATGATGATTTGGGACC 1174  
Db 336 TTGCGCTCCGTAAGAGGTACAGACTCGACGGCAACAGAGATGCTGGCCATGGGGTTC 395  
Qy 1175 ATGAACTGATTTGTTCTTTCACTCTTGTCTTGTCTTCACTCAACAGAGCAATTTTCGCTG 1234  
Db 396 TCCAACCTTGTGTTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455  
Qy 1235 GCTGTGAACATTAACGCTGGATGCAACACAGCAGCTTC 1272  
Db 456 GCAGTGAACCTTACGCGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 493

RESULT 10  
US-09-720-317A-9  
; Sequence 9, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR FILING DATE: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 9  
; LENGTH: 484  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; NAME/KEY: unsure  
; LOCATION: (398)  
; OTHER INFORMATION: n = A, C, G or T  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (483)  
; OTHER INFORMATION: n = A, C, G or T  
US-09-720-317A-9

Query Match 4.8%; Score 117.4; DB 4; Length 484;  
Best Local Similarity 55.7%; Pred. No. 2.5e-22;  
Matches 265; Conservative 0; Mismatches 207; Indels 4; Gaps 2;  
Qy 717 TTCTTGGCCCTTGAGCATTTTACCCATGGAGCTGATATCATATCAGTATGCGCTCTGTTT 776  
Db 2 TTACAGGCATCAAGAGCTTTTACAAAGAAACCGATATAATTTCCGTGATGAGCTGAAGTC 61  
Qy 777 TCACCCAAACTCATGA---GTGGAGGTGGAAAGTCTGTGTTAGTATGTTCTTCATTT 833  
Db 62 CAACAGAGGGCATACAGGTGGAAATGGCAAACTATTGTGATGGCATAACTTTCCTTG 121  
Qy 834 TCTTCTCTTAGCACAGATCTTCAAGCAAAACGACCAAGGTTTTTTTGGGTGTGAG 893  
Db 122 CATTCCTCTGCTTGCCAAAGTACATTGGAAAGAAAGTAAGGAAGTTCTTCTGGGTGCCAG 181  
Qy 894 CAATGGCGCATGAGCTCGTTATATTGGGAAGTCTCTTGGTTTATTTTCACTCAGCGCG 953  
Db 182 CTATTGCTCTATAACTTTCAGTTATTTTGGCAACCCCTTTTGTGTTTCATTACTCGTGTG 241  
Qy 954 AGAAGCACGCTGTGAAGTGTAGGAGAACTGAAGAGGGTTTTGAATCCACCATCACTCA 1013  
Db 242 ACAGCAAGGTGTTTCAAGTTGTTAAACCATCAAAGGGCATTAACCCATCATCACTAGTCC 301  
Qy 1014 CAAATCTGTTATTTGTGTCGCCCTTACATGACTACAGCTGTCAAAACCTGGCATTTGCTTG 1073  
Db 302 ACAAAATTTATTTCACTGGTCCATTTGTTGCAAAAGGTTTCAAGATCGGTGTCAATTCOG 361  
Qy 1074 GATCATATCATCTTGGGGAAGGAATAGCAGTAGGAAGAGAGCTTTTGCATGTATAAAATTT 1133  
Db 362 CCATGATCGGTTTAAACGGAAGCTGTGGCAATTTGGGANGAGCTTTTGTGCTCTGAAGACT 421  
Qy 1134 ACAATATTGATCGCAACAGAGATGATAGCTATTTCGGACCATGAACGTAGTTGGT 1189  
Db 422 ATCAATTAGATTGGAACAAGGAGATG-GAACACTTGGAACTATGAACATACAAAGGT 476

RESULT 11  
US-09-785-381-4  
; Sequence 4, Application US/09785381  
; Patent No. 6602992



```

US-09-785-381-2
; Sequence 2, Application US/09785381
; Patent No. 6602992
; GENERAL INFORMATION:
; APPLICANT: DALLOS, Peter
; APPLICANT: ZHENG, Jing
; APPLICANT: MADISON, Laird
; TITLE OF INVENTION: A MAMMALIAN PRESTIN
; FILE REFERENCE: 0290-37U1
; CURRENT APPLICATION NUMBER: US/09/785,381
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,461
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 4113
; TYPE: DNA
; ORGANISM: Meriones unguiculatus
US-09-785-381-2

```

Query Match	3.3%; Score 79.8; DB 4; Length 4113;
Best Local Similarity	46.1%; Pred. No. 2.3e-11;
Matches 304; Conservative 0; Mismatches 352; Indels 3; Gaps 1;	
Qy	1060 TGGCATTTGCTGGGCATCATATACATTTGCGGAGGAATAGCAGTAGAAGACGTTTGC 1119
Db	1249 TGCCATTTGCCATAGCCATCGTTGGATTTTCAGTGACAAATTTCCATGCGCCAAAACCTTGGC 1308
Qy	1120 AATGTATAAAAAATTACAAATTTGATGGCAACAAAGAGATGATAGCTATTGGGACCATGAA 1179
Db	1309 GAATAAGCATGGCTACAGGTTGATGGCAATCAGAGCTCATGCTTTGGGATATGCAA 1368
Qy	1180 CGTAGTTGGTTCTTTACCTCTTGCTACTCCTCACAAAGGACCAATTTTCGGTTTCGGCTGT 1239
Db	1369 CTCCATCGGATCTCTCTTCCAGACCTTCTCCATTTCTGTCTCTCGAGCCCTTGT 1428
Qy	1240 GAACATAACGCTGGATGCAAGACAGAGCTTCCACATTTATATGTGATCTTGCAGTAAT 1299
Db	1429 TCAGGAGGGAACTGGAGGGAAAAACACAGCTCGCAGGTTGCTTGGCCCTCGCTGATGTTCT 1488
Qy	1300 GTTGACATTTGTTATCTCTGACACCTTGTTCCTATACACTCCCTGGTGGTGCTATCAGC 1359
Db	1489 GCTGGTCAITTTTAGCCACTGGATTCTCTTTTGAGTCAATGGCCACAGCTGTGCTCTGGC 1548
Qy	1360 TATTATCGTATCTGCAATGCTTGGACTCATAGATTATGAAGACAGCCATCCATCTATTAA 1419
Db	1549 CATTTGATCGTGAACCTGAAAGGGATGTTTATGCAGTTCTCAGATCGCCCTTCTCTG 1608
Qy	1420 GGTT--GACAAATTTGACTTTTGGTGTGCGATGAGTGCAATACATTTGGCGTGTCTTTGG 1476
Db	1609 GAGAACCCAGCAAAATAGAGCTGACCACTGGCTTACCACCTTTGTGTCTCCCTGTTCCT 1668
Qy	1477 CAGTGTGAATTTGGCTTAGTCATAGTATTTGTAATCTGTACTTCGGGTACTTCTATT 1536
Db	1669 GGGCTTGGACTACGGACTGATTACTGCTGTGATCAITTTGCTCTGCTAGCTGATTTACAG 1728
Qy	1537 TATTGCAAGGCCAAGGACATTTGTTTTGGGCAACATTTCCAAATTTCTGTGATATACCGAAA 1596
Db	1729 AACCCAGTCCGAGCTACAAGTCTTGGGCGAGCTCCCTTGACACCGATGTTATACATTGA 1788
Qy	1597 TGTGTGACACTATCCAAATGCCAAACATGTTTCTCTGGAATGCTAAATCTAGAGATTGATGC 1656
Db	1789 CATAGACGCATATGAGGAGGTGAAGAAATTCCTTGGTAATAAAATATTCAGATAAACGC 1848
Qy	1657 ACCAATTTACTTTGCCAATCCGAGCTATTTAAGAGAAAGGATCACAGGTGGATGATG 1715
Db	1849 CCCAATTTACTATGCAACAGTCACTTTGTATAGCAACGCCCTTAAAAAGAAAGACTGTGTG 1907

RESULT 14  
US-09-248-796A-6643  
; Sequence 6643, Application US/09248796A

```

% Patent No. 67437137
%
% GENERAL INFORMATION
%
% APPLICANT: Keich
%
% TITLE OF INVENTION:
%
% TITLE OF INVENTION:
%
% FILE REFERENCE: 1
%
% CURRENT APPLICATION:
%
% CURRENT FILING DATE:
%
% PRIOR APPLICATION:
%
% PRIOR FILING DATE:
%
% PRIOR APPLICATION:
%
% PRIOR FILING DATE:
%
% NUMBER OF SEQ ID:
%
% SEQ ID NO 6643
%
% LENGTH: 2487
%
% TYPE: DNA
%
% ORGANISM: Candida
US-09-248-736A-6643

```

Query Match	3.2%	Score 77.6	DB 4	Length 2487
Best Local Similarity	48.8%	Pred. No. 7.2e-11		
Matches 209	Conservative 0	Mismatches 219	Indels 0	Gaps 0
QY	257	CTTCAGTTCCTCTCCCATTTTCGAATGGGCTCCCAATACACCTTTCAGTTCITGAAA	316	
Db	253	CTTATTTCATGTGTTCCAAATCTTAAATGGATTTTGCAATTACAAATATTAGATGGTTATAT	312	
QY	317	GCTGACCTCATAGCTGSCATCACCATCGCTAGCTTGCCCAATTCCTCAGGSCATCAGTTAT	376	
Db	313	GGTGACTGTGGTGTCTGGTATTACTGTGGGAGTTGTTTTAGTGCCTCAATCTATGTCATAT	372	
QY	377	GCCAAAGCTGCCAAACCTCCCTCCAAATCTTTGGACTATATTCGAGCTTTTATFACCACCATTG	436	
Db	373	GCCCAATTAGCTGGGTAGAAGCTCAATATGGTTTATATTCCTTTGTCGGTGTTTTC	432	
QY	437	ATTATACGATGATGGGTAGCTCGAGGATTTGGCAGTGGGGACTGTGGCGGTGGATCG	496	
Db	433	ATTATTATCATTTTTCGCCACCTCAAAGAATGTTTCCATCGGTCTCTGCTGTATTGTCTCC	492	
QY	497	CTTCTGATGGGTTGCATGTGTAGTATGCGGTGATGCCGTGATCCCAATGAAGACCCAAAGCTTAC	556	
Db	493	TTGCAAGTGCTAAGGTCATGTGCTCATGTTTCAAGATAAGTTTGGTGACAAATGCTGCT	552	
QY	557	CTCACCTGGCTTTCACAGCTACATTTATTTGCTGGTGTGTTTTTTCAGGCTGCTTTGGGTCG	616	
Db	553	CCTGAATGCCACATTTTGTCAATGATTTGTGGTGGTATCGCTCTTGGTATGGTTTG	612	
QY	617	TTTAGGTGGGGTTGATCGTGGATTTTCTGTACATGCAACCAATAATAGGGTTTCATGGGA	676	
Db	613	TTACGTTTAGGATTCATTTTAGAATTTATTTCTATCCAGCGGTGATGGGTTTCATGACT	672	
QY	677	GGAGCAGC	684	
Db	673	GGTCTGCG	680	

```

RESULT 15
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 03:14:21 ; Search time 2765 Seconds  
(without alignments)  
5795.308 Million cell updates/sec

Title: US-10-762-049-17

Perfect score: 2449

Sequence: 1 gcacgagctagctgcacat.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1521.6	62.1	2008	18	US-10-424-599-131116
2	1018.2	41.6	1977	9	US-09-938-842A-1502
3	1018.2	41.6	1977	11	US-09-938-842A-1502
4	846	34.5	2372	19	US-10-437-963-38184
5	790.2	32.3	1518	18	US-10-424-599-131117
6	677.4	27.7	2727	20	US-10-425-115-50601
7	670.4	27.4	4390	19	US-10-437-963-77012
					Sequence 131116,
					Sequence 1502, Ap
					Sequence 38184, A
					Sequence 131117,
					Sequence 50601, A
					Sequence 77012, A

8	658.6	26.9	2366	18	US-10-424-599-65103	Sequence 65103, A
9	650.6	26.6	2651	20	US-10-425-115-182207	Sequence 182207, A
10	647.2	26.4	2780	19	US-10-437-963-77005	Sequence 77005, A
11	600.6	24.5	2001	19	US-10-437-963-95853	Sequence 95853, A
12	571.4	23.3	2429	19	US-10-437-963-73410	Sequence 73410, A
13	567	23.2	2656	18	US-10-425-114-15719	Sequence 15719, A
14	567	23.2	2713	20	US-10-425-115-75763	Sequence 75763, A
15	520.2	21.2	536	18	US-10-424-599-32308	Sequence 32308, A
16	499.6	20.4	680	18	US-10-424-599-32307	Sequence 32307, A
17	489	20.0	2336	18	US-10-424-599-52175	Sequence 52175, A
18	482.2	19.7	3201	20	US-10-425-115-44188	Sequence 44188, A
19	477.4	19.5	1992	18	US-10-425-114-13333	Sequence 13333, A
20	465.2	19.0	2242	20	US-10-425-115-40196	Sequence 40196, A
21	433.8	17.7	585	18	US-10-424-599-108738	Sequence 108738, A
22	414.6	16.9	2251	19	US-10-437-963-72062	Sequence 72062, A
23	401.4	16.4	1980	17	US-10-260-238-1066	Sequence 1066, Ap
24	396.8	16.2	1809	18	US-10-424-599-83137	Sequence 83137, A
25	389.6	15.9	2506	18	US-10-424-599-74393	Sequence 74393, A
26	372.4	15.2	1303	18	US-10-424-599-139734	Sequence 139734, A
27	361.8	14.8	1247	18	US-10-424-599-130951	Sequence 130951, A
28	361.2	14.7	2156	18	US-10-424-599-60281	Sequence 60281, A
29	359.6	14.7	2156	20	US-10-739-930-3837	Sequence 3837, Ap
30	348.8	14.2	1551	18	US-10-424-599-54350	Sequence 54350, A
31	339.8	13.9	674	19	US-10-767-701-701	Sequence 701, App
32	329.4	13.5	1115	19	US-10-767-701-9435	Sequence 9435, Ap
33	323.8	13.2	3252	19	US-10-437-963-78357	Sequence 78357, A
34	315.4	12.9	1019	19	US-10-767-701-11661	Sequence 11661, A
35	311.6	12.7	1069	20	US-10-425-115-3619	Sequence 3619, Ap
36	298.6	12.2	2913	19	US-10-437-963-76453	Sequence 76453, A
37	298.2	12.2	1174	20	US-10-425-115-75765	Sequence 75765, A
38	293.2	12.0	2346	20	US-10-425-115-25539	Sequence 22539, A
39	291.4	11.9	1632	18	US-10-424-599-74389	Sequence 74389, A
40	285.6	11.7	565	18	US-10-424-599-62889	Sequence 62889, A
41	271.4	11.1	788	20	US-10-425-115-75767	Sequence 75767, A
42	258.8	10.6	1883	18	US-10-424-599-103813	Sequence 103813, A
43	255	10.4	459	18	US-10-424-599-120407	Sequence 120407, A
44	249.8	10.2	735	18	US-10-424-599-35721	Sequence 35721, A
45	249.4	10.2	841	18	US-10-424-599-135930	Sequence 135930, A

ALIGNMENTS

RESULT 1

US-10-424-599-131116  
; Sequence 131116, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 131116  
; LENGTH: 2008  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_89404C.1  
US-10-424-599-131116

Query Match 62.1%; Score 1521.6; DB 18; Length 2008;  
Best Local Similarity 94.9%; Pred. No. 0;  
Matches 1584; Conservative 0; Mismatches 84; Indels 2; Gaps 1;  
Qy 780 CCCAAACTCATGAGTGGAGTGGGAAGTCTGTGTAGGATGCTTCTTCATTTCTTCC 839  
Db 1 CCCAAACGCATGAGTGGAGTGGGAAGCGCTGTGTGGTGTCTGTATTTCTTCC 60





Db 118 GAAACTCTGTTTCAGACGACCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAA 177  
Qy 245 TTCAATGCTTGGCCCTTCAGTTCCTTCCCAATTTTCGAATGGGCTCCCAAAATACACCTTT 304  
Db 178 TTTGTGTTAGGCTCAAAATCTCTCTCCGATTTTCGAATGGGACCAACGCTACATCTC 237  
Qy 305 CAGTCTTGAAGCTGACCTCATAGCTGGCANTACCAATCGTGTAGCTTGGCCATTCCTCAG 364  
Db 238 AAGTCTTTCAAATCAGATCTCATCGCGGAATCACCATCGCTAGTCTCGCCATCCCTCAG 297  
Qy 365 GGCAATCAGTTATGCAAGCTGCGCAACTCCCTCCAATTCCTGGACATATTCGAGCTTT 424  
Db 298 GGCAATCAGTTACGCCAACTTGCTACTTGGCCCAATTTCTGGCCTTTATTCGAGTTT 357  
Qy 425 ATACCACCATTTGATTTATCGCATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGACTGTG 484  
Db 358 GTACCGCATTTGGTATACGCGGTCTAGGAGTTTCAAGGGACTTAGCGGTGGGAACGGTT 417  
Qy 485 GCGGTTGGATCGCTTCGTATGGGTTTCGATGTGAGTAAATCGCGTTGATCCCAATGAAGAC 544  
Db 418 GCGGTTGCGTCTCTGTTGACAGGTGCGATGCTGAGCAAGAAGTTGATGCTGAGAAGAT 477  
Qy 545 CCAAGCTTTACCTCCACCTGGCTTTCACAGCTACATTTATTTGCTGTGTT-TTTCAGGC 603  
Db 478 CCTAAGCTTTACCTTCACTTTCCTTTCACGCCACTTTTTCGCGCGGTTCTCGAAGCC 537  
Qy 604 TGCCTTGGGCTCTGTTTAGTTGGGTTTGATCGTGGATTTTCTGTCAATGCAACCAATAAT 663  
Db 538 TCTCTTGAATTTTCAGSGTTAGSGTTTCATAGTGGACTTTCTATCGCATCAACGATAGT 597  
Qy 664 AGGTTTCATGGGAGGACGACGCGGTGTTGATCGTGGATTTTCTGTCAATGCAACCAATAAT 723  
Db 598 AGGATTCATGGGAGGACGACGCGGTGTTGATCTGCAACAGCTTTAAGGCTATTTTCGG 657  
Qy 724 CCTTGACATTTACCCATCGAGCTGATATCATATCAGTCAATGCTGCTGTTTTCACCCA 783  
Db 658 ACTTAAACATTTACAGACTCTACCGATGTTAICTGTGATGCGTTCCGTTTCTCCCA 717  
Qy 784 AACTCATGAGTGGAGTGGGAAAGTGTGTGTTAGGATGTGTTTCAATTTTCTTCCTCCT 843  
Db 718 AACTCAGAGTGGAGATGGGAAAGTGGCGTTCTTGGATGTGTTTCTATTTCTTCTCT 777  
Qy 844 TAGCAAGAATATCTTACGAAAAAACAACAGGTTTTTTTGGGTGTCAGCAATGCGGCC 903  
Db 778 CTCACCCAGATATTGAGCATCAAGAAACCAAAATCTTTTGGTGGCGGATGGCTCC 837  
Qy 904 ATTGACGTCGTTATATTGGMAAGTCTCTGTTTATTTTCACTCAGCCGAGAGCAAGC 963  
Db 838 TTTGACCTCAGTGATCTTGGAAAGTCTCTTGGTTTACTTCACTCAGCGCTGAGAGACATGG 897  
Qy 964 TGTGAAAGTATAGGAGAACTGAAGAGGGTTTGAATCCACCATCACTCAAAATCTGGT 1023  
Db 898 TGTTCAGTATAGGGGACCTGAAGAAAGGTTTGAATCCACTCTCGGTTCTGATCTCAT 957  
Qy 1024 ATTTGTGTCGCTTACATGATCAGCTGTCAAACTGGCAATGTGCTGTTGGGATCATATC 1083  
Db 958 CTTTACTTCCCTTACATGTCCACAGCTGTCAAACTGGCCTCATCACTGGCATCAATTGC 1017  
Qy 1084 ACTTGGCGGAGGAATAGCAGTATGGAAGAAGCTTTGCAATGTATAAAATACAAATTTGA 1143  
Db 1018 TCTGCG-GAAGGAGTATGAGTGGGAGGATTTTGGCATGTTTCAAGAACTACAACTAG 1076  
Qy 1144 TGGCAACAAAGAGATGATAGTATTTGGGACCATGAACGTAGTTGGTTTCTTCCTCTTG 1203  
Db 1077 CGGGAACAAAGAGATGATAGGTTTGAATGATGAACATGTTGGTTCTTCACTCTTG 1136  
Qy 1204 CTACTCACAACAGGACCATTTTCGCGTTGCGGTGTGAATATAAAGCTGGATGCAAGAC 1263  
Db 1137 TTACTCTCACACCGGACCAITTTTCAAGTTCGCGAGTGAACCTAACACGCGGTTGCAAGAC 1196  
Qy 1264 AGCAGCTTCCAACTTATATGTCACCTTGCAGTAAATGTTGACATTTGTTTCTCTGACCC 1323

Db 1197 CGCAATGTCGAACATAGTAGTGGCGATTGGGGTTATGTTTCACACTCCTCTTCTCTCACACC 1256  
Qy 1324 CTTGTTCCATTTACACTCCCTGGTGGTCTATCAGCTATTTATCTATCTGCAATGCTTGG 1383  
Db 1257 GCTTTTTCATACACACCACTCGTCTCTCTCTGCAATCATCATATCGCAATGCTCGG 1316  
Qy 1384 ACTCATAGATATGAAGAGCCATCCATCTATTTAAGTTGACAAATTTGACATTTGGT 1443  
Db 1317 ACTCATTTGATATCAAGCTGCCATCCATCTCTGGAAGTTGACAAGTTTCGACTTCTCTCGT 1376  
Qy 1444 GTGCATGAGTGCAATATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTTAGTTCATAGC 1503  
Db 1377 CTGCATGAGGCGCTACGTTGGGTCGTATTCGCGAGTTGAGAGTTGGACTCTGCTGAGC 1436  
Qy 1504 TATTTGAAATATCTGTACTCTCGGTACTTCTATTTATTTGCAAGGCCAAGGACATTCGTTTT 1563  
Db 1437 GGTGGCGATATCTATAGCGAGGTTGTGCTGTTGTGTCGAGGCCAAAAAATCGCGTGAA 1496  
Qy 1564 GGGCAACATTTCCAAATTTCTGTGATATACCGAATGTTGACACTATCAAAATGCNAAACA 1623  
Db 1497 GGGAAACATACCAAAACAGCATGATCTATAGGAACACTGAGCAGTACCCATCATCAAGAAC 1556  
Qy 1624 TGTTCCTGGAATGCTAAATTTAGAGATTTGATGACCAAAATTTACTTTGCCAAATGCCAGCTA 1683  
Db 1557 CGTTCCTGGTATTTCTCATCTTGGAGATTGATGCTCCCATCTACTTTGCTAATGCCAGTTA 1616  
Qy 1684 TTTAAGAGAAAGGATCAAGGTGGATTTGATGAAGAAAGAAAGAAATTTAAAGCTACAGG 1743  
Db 1617 CTTGCGTGAAGAAATCATTAAGGTGGATTTGATGAAGAAAGAAAGAGAGTTTAAACAATCAGG 1676  
Qy 1744 GGNAGCTAGTTTGCAGTATGTTATTAATTTGATATGATGCTGTTTGGAAACATTTATACAAG 1803  
Db 1677 AGAGACGACCTTACAATATATTATCTCGATATGTCAGCTGTTGGTAAATTCGACACAAG 1736  
Qy 1804 TGGAAATAGTATGCTTTGAAGAGGTGAAGAGATTTACAGAGAGAAAGAGAGCTACAGCTTGT 1863  
Db 1737 CGGTATTGATGATGCTGGAAATTTAAGAAAGTCTTGAAGAGAGCGTTAAAGTTGGT 1796  
Qy 1864 TTTGGTCAATCTCTGTTAAGTGAAGTGAAGAAACTGAACAAATCGAAGT--TCCAAA 1920  
Db 1797 ATTGTCAAAATCCAAAAGGAGAGGTCGTGAAGAAATTTAACAGATCCAAATTCATCGGTGA 1856  
Qy 1921 TCATTTAGGAGAAATGATCTATCTGACTGTTTGAAGAGCGGTTGGAGCACTT 1980  
Db 1857 TCATTTGGGCAAGAGTGGATGTTCTTAACGGTAGGAGAGCAGTGGAGGCTTTGAGCTA 1916  
Qy 1981 CAATCTACGTGCAAGCAAAACCAACCC 2007  
Db 1917 CATGCTTCACAGCTTTTAAACCCGAACC 1943

RESULT 3

US-09-938-842A-1502  
; Sequence 1502, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRI1300-3  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1502

[illegible]





Db 300 GGGTACAAAGTCGGCGTCCCAACCCGAGAGAAACCTCTCTCGCCGAGATCTCTGACGCGGTG 359  
Qy 182 AAGGAGAGCTTCTTCCCTGTAGTACCTTTGAGGAGAGTTCAGAAACAAGACAGCTTCCAAG 241  
Db 360 AAGGAGAGCTTCTTCCGCGAGACCCGCTGGGAGTACAGGACACAGCCAGCTCCAAG 419  
Qy 242 AAGTTCATGCTTGGCTTTCAGTCTTCTTCCCAATTTTCGAATGGGCTCCCAATAAC 301  
Db 420 AAGATCTGGCTCGGCTCGAGCAATCTTCCGCGTCTCGAATGGAGCAGCAATTAATCTCC 479  
Qy 302 TTTTCAGTCTTGAAGCTGACCTCATAGCTGGCATCACCTCGTAGCTTGGCCATTCCT 361  
Db 480 CTCCGCAAGTTCAGGGGTGACTTCACTCGTGGGTCCACCATTGCAGCCTCTGCATACCC 539  
Qy 362 CAGGCGATCAGTATGCAAGCTCGCAACCTCCCTCCCAATTTCTTGGACTATATTCGAGC 421  
Db 540 CAGGACATCGGTATTTCTAAGCTTGTCTAACTTGCAGCAGAGGTGGGACTATACAGTAGC 599  
Qy 422 TTTATACACCATTTGATTTATGCAATGATGGTGTAGCTCGAGGGATTTGGCAGTGGGACT 481  
Db 600 TTTGTCGCGCTCTGATATAGCAGATGATGGGAGCTCCAGGACATAGCCATCGGTCCA 659  
Qy 482 GTGCGGTGCGATCGCTTCTGATGGGTTCGATGTTGAGTATGATCCGTTGATCCCAATGAA 541  
Db 660 GTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719  
Qy 542 GACCCAAAGCTTTTACCTCCACCTGGCTTTCACAGCTACATTTATGCTGGTGTCTTTCAG 601  
Db 720 CACCCGCTCGAGTACAGGCGCTAGCCTTTCACAGCAACCTTTTCGAGGAGTCACTCAG 779  
Qy 602 GTCGCTTGGGTCTGTTTGTAGTGGGTGATGCTGATGTTTCTGTCACATGCAACCAATA 661  
Db 780 GCAGCCCTCGGATCTTTCAGGCTAGGCTTTATCATAGAGTCTTGTCTCATGCTGCGCAT 839  
Qy 662 ATAGGTTTCATGGAGGAGCAGCACCGTGGTGTCTGAGCAACCTAAATCGATTCCT 721  
Db 840 GTCGGAATTCATGGCGCGCTGCCATCACCATTGCGCTTCAGCAGCTGGAAGGATTCCT 899  
Qy 722 GGCCTTGAGCATTTTCAACCCATGGAGCTCATATCATATCAGTGTGCTGCTTTCACC 781  
Db 900 GGAATTCGAATTTTCAACCAAGAAATCCGATATGATCTGTTATGAATCAGTTGGGA 959  
Qy 782 CAAACT---CATGAGTGGAGTGGAAAGTGTGTTGATGATGTTGTTCTTCAATTTCTTC 838  
Db 960 AATGTTCAACCATGGGTGGAATTTGGCAGCAATACTGATAGGAGCGACCTTCTGGCATTC 1019  
Qy 839 CTCTTAGCACAAGATCTTACGAAACCAAGCAAGTCTTTCGGGTGTCAGCAATG 898  
Db 1020 CTTCTGTTTCCAGTACATTTGGGAAAGGAATTAAGAAAGCTTCTTGGGTGTCGCAATC 1079  
Qy 899 GCGCATTTGAGCTCGCTTATTTGGGAAGTCTCTTGGTGTATTTTCACTCACGCGGAGAAG 958  
Db 1080 GCACCTCTACTTCTGGTATCATATCCATATTTTGTGTATCATCTCTGTGAGATAG 1139  
Qy 959 CAGGCTTTGAAGTATAGAGAACTGAAGAGGGTTTGAATCCACCATCACTCACAAAAT 1018  
Db 1140 CATGGCTTGCATTTGTCAAGAAATTAAGGAAGGCAATCAACCCACCTTCAGCTAGCTA 1199  
Qy 1019 CTGGTATTTGTGCTTACATGACTACAGCTGTCAAACTGGCATTGCTGGCATC 1078  
Db 1200 ATATACTTCCCGGCCCATCTTGGCCAGCAGGATTCAAAATTTGGGATAGTAGTGGAAATG 1259  
Qy 1079 ATATCACTTCGGGAAGGAATAGCAGTAGTAGGAAGAGCTTTTGCATGTATAAAATTAACAT 1138  
Db 1260 ATAGGCTTAACGGAGCGATTTGCAATTTGGAAGAACATTTTCAGCCCTCAGAGTATCCGG 1319  
Qy 1139 ATTGATCGCAACAAGAGATGATAGCTATTGGGACCAATGAACGATGTTCTTTCACC 1198  
Db 1320 ATAGATGGGAACAAGAAATGATGGCTCTAGGGACCAATGAACATTTGTTGGTTCACTGACT 1379  
Qy 1199 TCTTGCTACCTCACACAGGACCAATTTTCGGGTTCGGCTGTGAATTAACGCTGGATGC 1258

Db 1380 TCTTGCTACGTAGCCACAGGTTCTTTCTCGCGTCAAGCTTAATTAATGCTGGCTGC 1439  
Qy 1259 AAGACAGCAGCTTCCAAACATTAAATGTCACTTGCAGTAATGTGTGACATTTGTTATCTCTG 1318  
Db 1440 AAAACAGCAGTGTCAACCGTTGTTATGTCATTTGCTAATTTGCTAATGCTTACTTTGCTACTGATC 1499  
Qy 1319 ACACCTTGTTCATTACACTCCCTCGTGGTGGTGTATCAGCTATTATTCGTATTCGAAATG 1378  
Db 1500 ACCCCATTTGTTCAAGTACACTCCAAATGCCATCTTTCTTCGATCATCATCAGCAGTG 1559  
Qy 1379 CTTGGACTCATACATTAATGAAGCAGCCATCCATCTATTTAAAGGTTGACAAATTTGACTTTT 1438  
Db 1560 CTTGGTTTAAATGCACTATGAATCGGCTTACCTTATCTGGAAGTTGACAACTGGACTTT 1619  
Qy 1439 GTGGTGTGATGATGATCATACTTGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTC 1498  
Db 1620 CTAGCATGATGGAGCAATTTTGGAGTCAATTTTTCATCGGTGGAGTATGGCTTGCCTC 1679  
Qy 1499 ATAGCTATTGTTAATATCTGTACTTGGGTACTTCTATTTATTCGAAGGCCAAGACATTC 1558  
Db 1680 AATTGCGGTTGCAATATCTCTTGTCTAAATTTCTTCCAAAGTAAACCGGCCAAGAACAGTT 1739  
Qy 1559 GTTTTGGGCAACATTTCCAAATTTCTGTATATACCGAAATCTTGCAGCACTATCAAAATGCA 1618  
Db 1740 TTACTTGGAAACCTTCCAGCAACATATATACGGAATGTAGAACAGTATCTCTGATGCT 1799  
Qy 1619 AAAACATGTTCTCGAATGCTAAATTTCTAGAGATTTGATGCACCAATTTACTTTGGCAATGCC 1678  
Db 1800 ACCAAGTTTCCAGGGTGTCTGATTTGTAGTGTGACTCAGCTATATCTTTCACAACTCT 1859  
Qy 1679 AGCTATTTAAGAGAAAGGATCAACAGGTGATTTGATGAAGAGAAAGAAATTTAAAGCT 1738  
Db 1860 AACTATGTTTAAAGAGAGAAATCTCTGAGTGGCTAAGAGATGAGGAGGAGCAACACAGGAC 1919  
Qy 1739 ACAGGGGAGACTGATTTGCGAGTATTTAATTTGATATGATGCTGTGGAACATTTGAT 1798  
Db 1920 CAGAAGTTTAAACAAACTGATTTCTAATTTGTTGACTGCTCTCTGTTAATTTGATTCGAC 1979  
Qy 1799 ACAAGTGGAAATAGTATGCTTGAAGAGGTGAAGAAATTTACAGAGAGAAGAGCTACAG 1858  
Db 1980 ACAAGTGGAAATCCATGCTTTGGAGGAGTTGGCGAAAGCTTCTGAAACACGCAAAATTCAG 2039  
Qy 1859 CTTGTTTTGGTCAATCTCTGTAGTGAAGTGAAGAGAACTGAACAAATCGAAGTTCCAA 1918  
Db 2040 CTGGTTCTTACCAATCCCGGCGCGGTGATCCAGAAGCTCCGCTCAGCGAAATTCAGC 2099  
Qy 1919 AATCATTTAGGGAAGAAATGGAATCTATCTGACTGTTTGAAGAGGCCGT 1965  
Db 2100 GACATGATTGGTGAAGAACAACTATTTCTCCTCACGGTCGGGACGCTGT 2146

## RESULT 7

US-10-437-963-77012  
; Sequence 77012, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 77012  
; LENGTH: 4390  
; TYPE: DNA

; ORGANISM: Oryza sativa			
; FEATURE:			
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76952C.1			
US-10-437-963-77012			
Query Match 27.4%; Score 670.4; DB 19; Length 4390;			
Best Local Similarity 60.3%; Pred. No. 3.7e-169;			
Matches 1109; Conservative 0; Mismatches 731; Indels 0; Gaps 0;			
QY	126	ACCAAGTCGAGGTTCCACCGCCACAGCGGTTTTCAAGTCTCTAAAGTACTCTTCAAGG	185
DB	504	ACAAATGTCGGCGCGCCGCAAGAAGAACCTCTCTGGCGGAGTTGCGCGGCAAGG	563
QY	186	AGACTTTCTTCCCTGTATGACCCCTTTGAGGCAGTTTCAAGAACACAGAGCTTCAAGAAAT	245
DB	564	AGAGCTTCTTCTCGAGCGACCGATGCGGGGTACAGGACACGCGAGTCCAGGAGC	623
QY	246	TCATGCTTGGCCCTTCAGTTCTTCTTCCCAATTTTGGAAATGGGCTCCAAATACACCTTTC	305
DB	624	TATGGCTCGCTTGCAGCAGCTCTTCCCGGTGTTGCAATGGGGCAGACAATACACCCCTCG	683
QY	306	AGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATGCTAGCTTGGCCATTCCTCAGG	365
DB	684	CCAAGTTCAAGGGCGACCTCAITGGCGGCTCACCCTTGCACGCCCTCGTCATACCTCAGG	743
QY	366	GCATCAGTTATGCAAGCTCGCCAAACCTCCCTCCAATTTCTTGACTATATTTGAGCTTTA	425
DB	744	ACATCGCTACGGAGCTTGCTTAACCTGCCACAGAGATTGGCTGACAGTAGCTTCG	803
QY	426	TACACCATTTGATTTATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACGTGTGG	485
DB	804	TCCCGCATTTGATATACGCTCTGATGGGCACTCAAGGGAGCTAGCAATGGGTCACAGTGG	863
QY	486	CGGTTGATCGCTTCTGATGGGTTGATGTTGAGTAAATGCGGTTGATCCCAATGAGACC	545
DB	864	CCGTCACTCACCTGCTTGGTACTCTCTCCAGGAGGAGATTGACTCAAGAAGAACC	923
QY	546	CAAAAGCTTTACCTCCACCTCGCTTTTCAAGCTACATTAATTTGCTGGTGTTTTTCAGGCTG	605
DB	924	CGCTAGATTACAGCGGCTCGCCCTTACAGCGACCTTCTTTGAGAGTCAACAGCGGG	983
QY	606	CCTTGGGTCGTGTAGTTGGGTTGATTCGTGATTTTCTGTCATGCAATGCAATATATAG	665
DB	984	CGCTGGGTTTCTCAGGCTAGGGTTTCAATCATAGCGTTCTCTGTCTCATGTGCCATCATCG	1043
QY	666	GGTTTCATGGGAGGAGCAGCAGCGTGTGTGCTGTCAGCAACATAAATCGATTCTTGGCC	725
DB	1044	GATTCATGGCGGCGCGCCATCACCATTTGCTCTTTCAGCAGCTTAAAGGCTTCCITGGAA	1103
QY	726	TTGAGCATTTTACCCATGGAGCTGATATCATATCAGTGATGCGCTCTGTTTTCACCCAAA	785
DB	1104	TTGCCAACTTTCACCAAGAAGACTGACATCATCTCGGTCTATGAATCAGTCTGGGAAATG	1163
QY	786	CTCATGATGGAGGTGGGAAAGTGTGTGTGTAGGATGTGTCTCATATTTCTTCCCTCTTA	845
DB	1164	TTCCATATGGGGCAATGGAACTGGCAGACAAATATTCATCGGAGCATCATTTTGGCATTC	1223
QY	846	GCACAGATATTTACGCAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGCCAT	905
DB	1224	CTCCTGGTTGCCAAGGCAAGAAGAACAAAGAGCTCTTCTGGGTCCAGCAATTTGCACCAC	1283
QY	906	TGACGTCCTGTATATTTGGGAAGTCTCTTGGTTTTATTTTCACTACGCGCGAGAAGCAGGTG	965
DB	1284	TCATTTGGGTGATCATTTCAACTTTGTTGCTCTACATCACTCGTGTGACAAACAGGCG	1343
QY	966	TTGAAATGATAGGAGAACTGAAGAAGGTTTTGAATCCACATCACTCACAAAATCTGGTAT	1025
DB	1344	TCGCAATTCGTAATAAAGCGTCAAGAAAAGGATCAATCCACCTTCAGCTAGCTGATATTT	1403
QY	1026	TTGTGTGCGCTTACATGACTTACAGCTGTCAAAATGTCATGCTGCTTGGCATCATATCAC	1085
DB	1404	TCACTGGCCCATACTTCTCAAAGGATTCAAAATTTGAAGTAGTAGCTGGAATGATAGCC	1463

QY	1086	TTTCGGAAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATAAAAATTAACAATTTGATG	1145
DB	1464	TTACCGAAGCGATTCAGTTTGAAGAACATTTTCTGCTGATTTGAACGATTACAGATAGATG	1523
QY	1146	GCAACAAAGAGATAGAGCTATTTGGGACCAATGAAAGTAGTTGCTTTTCACTCTTGTCT	1205
DB	1524	GGAACAAAGAAATTTGGCTCTAGGAACCAATGAATGTGCTGCTTCAATGACGCTTGTCT	1583
QY	1206	ACCTCAACACAGCACCATTTTTCGGCTGTGAACCTATTAACGCTGATGCAAGACAG	1265
DB	1584	ATATAGCCACAGGTGTTTTTGCACATCAGCAGTCAATTTGCTGCTGGCTGGAGTTAAACAC	1643
QY	1266	CAGCTTCCAACTATAATATGTCATTTGCAAGTAAATTTGACATTTGTTTCTTCCGACCCCT	1325
DB	1644	CAATGTCCAAATTTGTTATGCACTGTAGTATTTGCTTGCATCTCTGTGGTCACTCCAT	1703
QY	1326	TGTTCCATTTACACTCCCTGGTGGTCTATCAGCTATTTATCGTATCTGAAATGCTTGGAC	1385
DB	1704	TGTTCAAGTACACGCCCAATGCCCAATTTCTTCCATCATATCATCAGCAGTGTCTTGGCC	1763
QY	1386	TCATAGATTTATGAAGCAGCCATCCATCTATTTTAAAGTTGACAAATTTTGAATTTTGGTCT	1445
DB	1764	TATTTGACTTTGAATCAGCTACCTTTATCTGGAAGTTTGAAGTTGGACTTCAATGGCGT	1823
QY	1446	GCATGAGTCATACATTTGGCGTGGTCTTTTGGCAGTGTGAAATTTGGCTTTAGTCATAGCTA	1505
DB	1824	GCTTGGGGCAATTCCTTGGAGTAATATTTTTCATCTGTGGAGTATGGCTTCTCATTTGCGG	1883
QY	1506	TTGTAATATCTGTACTTTGGGTAATTTTATTTATTTTGAAGGCCAAGGACATTTCTGTTTGG	1565
DB	1884	TTGTAATATCACTAATAAAGTTCTGCTCCATGTAAACACGCGCCAAAGGACAGCTTTACTTG	1943
QY	1566	GCAACATTTCCAAATTTCTGTGATATACCGAAATTTTGAGCAGTATCAAAATGCAAAACATG	1625
DB	1944	GCAACTTTCCAAAGACGATTTATCTATAGAAATGTTGAACATATCCAGAGCTACCAAGG	2003
QY	1626	TTCTCTGGAATGCTAATTTCTAGAGATTGATGCAACCAATTTTACTTTGCAATGCCAGCTATT	1685
DB	2004	TGCCAGGATGCTAATTTGTAAGAGTGGACTCAGCAATATATCTTCAAACTCCAATTATG	2063
QY	1686	TGAAGAAAGGATCACAAGGTGATGTAAGAAAGAAAGAAATTAAGCTACAGGGG	1745
DB	2064	TTAAAGAAAGAAATGCTGAGATGGCTGAGAGATGAGAAAGAAACATCAAAAGGAAACAGAAT	2123
QY	1746	AGACTAGTTTGCAGTATGTTAATTTGATATGAGTGTCTTGGAAACATTTGATACAAAGTG	1805
DB	2124	TACCAAAATTTGAGTTTCTGATTTGACCTATCTCTGTAATGATATTTGACAAGTG	2183
QY	1806	GAAATAGTATGCTTGAAGAGGTGAAGAAATTTACAGAGAGAAAGAGCTTACAGCTTGT	1865
DB	2184	GAAATCCATGCTTCAAGAGATTTGTTGAGGACACTTTGAAAAGCGCCAGATTCAGCTGATTT	2243
QY	1866	TGGTCAATCTCTAAGTGAAGTGAAGAAATCTGAACAAATCGAAGTTCCAAATCATTT	1925
DB	2244	TGCGCAATCTCTGGCGGCTGTGATCCAAAAGCTCCGCTCAGCAAAATTCACAGAGCTCA	2303
QY	1926	TAGGGAAGAAATGGAATCTATCTGACTGTTTGAAGAGGCGCT	1965
DB	2304	TTGGTGAAGAAAGATATGCCCTGACAGTTGGTGGACGCCGT	2343

RESULT 8  
US-10-424-599-65103  
; Sequence 65103, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B



; CURRENT APPLICATION NUMBER: US/10/424,599									
; CURRENT FILING DATE: 2003-04-28									
; NUMBER OF SEQ ID NOS: 285684									
; SEQ ID NO 65103									
; LENGTH: 2366									
; TYPE: DNA									
; ORGANISM: Glycine max									
; FEATURE:									
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29800C.1									
US-10-424-599-65103									
Query Match 26.9%; Score 658.6; DB 18; Length 2366;									
Best Local Similarity 60.5%; Pred. No. 3.9e-166;									
Matches 1120; Conservative 0; Mismatches 724; Indels 7; Gaps 2;									
Qy	77	AGTGTAGATTATGAGTACCCCTTTGGGCATGAACAACTTTTGAGAGAGTGCACCAAGTCGAG	136						
Db	132	AATGCATGTACTATGCATTCTCATTTGCATAGAGATGTCATGGAAGTGCACCAAGTTGTG	191						
Qy	137	GTTCACCGCCACAGCGTTTTTCAAGTCTCTAAAGTACTCTTTTGAAGGAGACTTTCTTC	196						
Db	192	CCACCACCTCACAGAGCACCCCTTCAGAACTCAAGGGTAGGCTCAAGGAACTTTCTTC	251						
Qy	197	CCTGATGACCTTTGAGGCGAGTTCAAGAACCAAGCCAGCTTCCCAAGAGTTCTATGCTGGC	256						
Db	252	CCTGATGATCCTCTGCGCCAAATTCAGGGACAACCTCTTAAGAGAAAACCTGATCCCTTGA	311						
Qy	257	CTTCAGTCTTCTTCCCAATTTTCCGAATGGGCTCCCAATACACCTTTTCAAGTTCTTGAA	316						
Db	312	GCTCAATATGTGTTCCCTATTTCTCAATGGGCTCTAAGTATATCTCAAACTCTTCMA	371						
Qy	317	GCTGACCTCATAGTGGCATCACATCGTAGTGTGGCCATTCCTCAGGGCATCAAGTTAT	376						
Db	372	TCTGACCTTGTCTGGCCCTCACTATTGCTAGCTTGACATCCCGCAGGGAATGAGTTAT	431						
Qy	377	GCCAACTGCCAACCTCCCTCCAAATCTTGGACTATATCGAGCTTATACCAACCATTCG	436						
Db	432	GCTTAGCTTCCAAAGTCTTCCCTCAATTTTAGGACTTTATTTCTAGTTTGTGTCACCACT	491						
Qy	437	ATTATATGCGATGATGGGTAGCTCGAGGATTTTGGCAGTGGGACTGTGGCGTTGGATCG	496						
Db	492	GTCTATGCTGTTCTGGAACTCAAGGACCTTGGAGTGGACTGTGTTCTATTGCTTCT	551						
Qy	497	CTTCTGATGGGTTCGATGTTGAGTAATGCCGTTGATCCCAATGAAGACCCAAAGCTTTAC	556						
Db	552	CTTGTGATGGGATCCATGTTGCAATCAGGAAGTGTCTCCCAACACAGATCCAAATCTGTTT	611						
Qy	557	CTCCACCTGGCTTTCACAGCTACATTTATTTGCTGGTGTGTTTTTCAGGCTGCTGGTCTG	616						
Db	612	CTTCAGCTAGCTTTCACTTCAACATTTATTTGCGGCTCTTTTCAAGCTTTGCTTGAATC	671						
Qy	617	TTTAGGTTGGGGTTGATCGTGGATTTTCTGTCATGCAACCAACCAATATAGGGTTTCATGGGA	676						
Db	672	CTAAGGCTAGGCTTCATAATTGATTTTCTATCTAAGGCCATTTCTATTGGGTTTCATGGCT	731						
Qy	677	GGAGCAGCCACGGTGGTGTCTGCGACAACTAAATCGAATCTTGGCCCTTGAGCAATTC	736						
Db	732	GGAGCTGCTATTATTGTCTCACTGCAACAGCTCAAGAGCCGTGCTGGAATCACACATTC	791						
Qy	737	ACCCATGGAGCTGATATCATATCAGTATGCGCTCTGTTTTCACCCAACTCATGATGG	796						
Db	792	ACTAATCAGATGGGTCTGATTCCTGTTATGACTTCTGTTTTTCACAATATACATGATGG	851						
Qy	797	AGGTGGGAAAGTGTGTGTTAGGATGTGTTCTCATTTTCTCTCTCTTAGCACAAGATAC	856						
Db	852	TCATGGCAACAATATTGATGGGATTTGCTTCTTGCTACTACTACTATTAGCAAGAC	911						
Qy	857	TTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTCAGCAATGGCGCAATGACGTCGGTT	916						
Db	912	GTTAGCATTAAGGAAACCAAACTATTCTGGGCTCTCAGCTGGAGCTCTCTTATGTGTGTC	971						
Qy	917	ATATTGGGAATGCTCTTGGTTTTATTTCACCTCACGCCGAGAACGCTGTTGAGTGATA	976						

RESULT 9

US-10-425-115-182207  
; Sequence 182207, Application US/10425115  
; Publication No. US2004021427A1  
; GENERAL INFORMATION: Thomas J.  
; APPLICANT: La Rosa, Thomas J.



; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 182207  
; LENGTH: 2651  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_97750C.1  
US-10-425-115-182207

Query Match 26.6%; Score 650.6; DB 20; Length 2651;  
Best Local Similarity 60.2%; Pred. No. 6e-164;  
Matches 1117; Conservative 0; Mismatches 729; Indels 9; Gaps 2;

QY	121	AGTGCACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTT	180
DB	517	AGTGCAACAGGTGCGCGCGCGCGCGGAGCACCGCGAGCAAGATGAAGGTGAGGGT	576
QY	181	GAAGGAGACTTTCTTCCCTGATGACCCCTTTGAGGCAGTTTCAAGAACAAAGCCAGCTTCCAA	240
DB	577	GAAGGAGACTTTCTTCCCGACGACCGTTCCGGCGGTTCAAGGGGAGCGCGCGGGGAC	636
QY	241	GAAGTTCAAGCTGGGCTTCAGTTCTTCTTCCCAATTTTCCGAATGGCTCCCAAAATACAC	300
DB	637	GCAGTGGCTCATGGCGTCAGGTACCTCTTCCCAATCTCTGGACTGGGTGCGAGCTACTC	696
QY	301	CTTTGAGTTCTGGAAGCTGACCTCATAGTGGCATCACCATCGCTAGCTTGGCCATTCC	360
DB	697	CTTGTGCTCTTCAAGTCGACCTGCTGCGCGGCTCACCAATGGCCAGCGCTGCGCAATTC	756
QY	361	TCAGGGCATCAGTTATGTCGAAGCTGCGCAACCTCCCTCCCAATTTCTTGGACTATATTCGAG	420
DB	757	TCAGGGCATTAGCTACGCGAAGCTGCGAAGCTTGGCTCCCAATTTCTGGGCTGTATTCGAG	816
QY	421	CTTTATACCAACATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT	480
DB	817	CTTTCGTCGCGCGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	876
QY	481	TGTGGCGGTTGGATCGCTTCTGATGGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATG	540
DB	877	GGTGTGATCTGCTGCTGATCATGCGGCTCCATGCTGCGGAGCGCTGAGCCCGCCACCGC	936
QY	541	AGACCCAAAGCTTTTACCTCCACCTGGCTTTCACAGCTTACATTTATTTGCTGGTGTTCAT	600
DB	937	GGAGCGACGCTGTTCTGAGCTGGCTTTCACCTCCACCTGTTTCGCGGGGCTGTGTGA	996
QY	601	GGTGGCTGGGCTGCTGTTTGGGTTGATCGTGGATTTTCTGTCAATGCAACCAAT	660
DB	997	GGCTTCCCTGGGCATCTCAGGCTCGCTTCTGCTCATGACTTCTCTGTCAAAGCGAGCT	1056
QY	661	AATAGGGTTCAATGGAGGAGCAGCAGCGGTGGTGTCTGCAGCACTAAATCGATTCT	720
DB	1057	GGTGGGGTTTCAATGGCGGCGCGCCATCATCTGTTGGCGCTGCAGCAGCTCAAGGCGCTGT	1116
QY	721	TGGCTTTGAGCATTTACCCCATCGAGCTGATATCATATCAGTATGATGCTGCTGTTTTCAG	780
DB	1117	GGGCATCGTCCAATTACCAACCGAGATGGGCATCGTCCAGTATGAGCTCGCTCTTCCA	1176
QY	781	CCAAATCATAGTGGAGGTGGAAAGTGTGTTGTTAGGATGTTCTCAATTTCTTCCT	840
DB	1177	CCACACGAGCGTGTGCTGGGAGACGATCTCTATGGGCGCTGCTGCTCTCTCTCT	1236
QY	841	CTTTAGCACAGATCTTACGCAAAACGACCAAGTTTTTTTGGTGTCTGACCAATGGC	900
DB	1237	GCTGTGCGGCGAGCATGTGAGCATCAGATGCGCAAGCTTTTCTGGGTTTCGCGGTGCGC	1296

QY	901	GCCATTGACGTCGCTTATATTGGGAAGTCTCTTGGTTTATTTTCACTCACGCCGAGAGCA	960
DB	1297	GCCCCGTGGCATCGGTCAACCATCTCGACGCTGCTTGTCTTCTCTTCAAGCTCAGAACCA	1356
QY	961	CGGTGTTGAAGTATAGGAGAACTGAAGAAAGGTTTGAATCCACCACTACTCAAAATCT	1020
DB	1357	TGGCATCAGCATCATTTGGGCAGCTCAAGTGGCGGCTGAATCGCCCTCGTGGGACAAGCT	1416
QY	1021	GGTATTTGTTGTCGCTTACATCCTACAGCTGTCAAACTGGCATTTGTTGGGATCAT	1080
DB	1417	CCTGTTTGACACGGCGTATTTAGGCGCTCACCATGAAGACTGGCCTTGTCAACGGATCAT	1476
QY	1081	ATCACCTTGGGAAAGAAATAGCAATAGCAAGCTTTGCAATGTATAAAATTTACAAAT	1140
DB	1477	CTACCTGACGGAAGGATAGCGTTGGTAGAACATTTGGCTCCTCAAGGACTACCGAT	1536
QY	1141	TGATGCAACAAAGAGATAGTATTTGGGACCAATGAACGTTAGTTGGTCTTCTTCACTC	1200
DB	1537	AGATGGAAACAAAGGAGATGATGGCATAGGTTGATGAATGTTGTTGGGCTCTGCACATC	1596
QY	1201	TTGCTACCTCAACAGGACCAATTTTCGGTTTCGGCTGTGAACATAACGCTGGATGCA	1260
DB	1597	ATGCTACGTTAAACAGAGTGGCTTCTCCGCTCTGCTGTAAACCAACACGCGGCTGCA	1656
QY	1261	GACGAGCTTCCAAACATTAATGTCACTTGCAGTAATGTTGACATTTGTTATTTCTGAC	1320
DB	1657	GACTGCCATGTCACCGTGATCATGGCGTGAATGATGATGATGATGATGATGATGATGAT	1716
QY	1321	ACCTTGTTCATATACATCCCTCGTGGTGTATCAGCTATATATGATATGATATGATATG	1380
DB	1717	GCCACTGTTTGTGTATACACCAACGTTGCTCGGAGCGATCATCATCGCGCGGTGAT	1776
QY	1381	TGACCTCATAGATATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTCT	1440
DB	1777	CGGCTGATGATTTCCCGCGGTGTACCACTCTGGAAGATGGACAGATGATGATTTCT	1836
QY	1441	GGTGTGATGATGATATATTTGGGCTGCTTTCGGCAGTGTGAAATTTGGCTTAGTAT	1500
DB	1837	GGTGTGGTTTGGCGTTTGGCGGCTCATCTTCATCTCAGTCCAAAGAGGCTTTCGAT	1896
QY	1501	AGCTATTTGATATCTGTACTTTCGGGTACTTCTATTTATTTGCAAGGCCAAGACATTC	1560
DB	1897	AGCGGTGTTGATATCTATATTTAGGTTGTTGATGAGATCACGAGGCCCAAGATGATG	1956
QY	1561	TTTGGGCAACATTCAAATTTCTGTATATACCGAAATGTTGAGCACTACAAATGCAAA	1620
DB	1957	TCAGGGCAACATCAAGGGGACTGATATTTACAGACCTGCACTACAGAGAGGCCCA	2016
QY	1621	ACATGTTCTGGAATGCTAAATTTAGAGATTTGATGACCAATTTTACTTTTGCCTTAC	1680
DB	2017	AAGAGTTTCTGGGTTCTTGATCTTGGCCATTGAAGCACCGATAACTTCGCCAATCCA	2076
QY	1681	CTATTTAAGAGAAAGATCAAGGTGATTTGAAGAAAGAAAGAAAGAAATTTAAAGCTAC	1740
DB	2077	CTACCTGATGAAGAGATTTAAAGATGGAT-----AGAGGAAGAAATCTTTTGAACAG	2130
QY	1741	AGGGGAGACTAGTTTGCAGTATGTTATAATGATATGATGCTGTTGGAAACATTTGATAC	1800
DB	2131	TAAACATCTGAATCCATTTCAATCTTGGATCTGTGAGCTGTTCTGCAATTTGACAC	2190
QY	1801	AAAGTGAATAAGTATGCTTTGAAGAGGTGAAGAGATTACAGAGAGAAGAGACTACAGCT	1860
DB	2191	AAAGTGGCTAGCGTTTCTCTCATTTGACATAAAGAAATCAATAGAGAAACGTTGCTGGAG	2250
QY	1861	TGTTTTTGTCAATCTGTAAAGTGAAGTGAAGAACTGAACAAATCGA---AGTTCCA	1917
DB	2251	TGTGCTTGTCAATCCAACTGGAGAGTCAATGAGAGAAATACAACTGCAACGAGGCTGA	2310
QY	1918	AAATCATTTAGGGAAGAAATGATCTATCTGACTGTTGAAGAGGCGGTTTGGAGCA	1972
DB	2311	AAACTATTTTAGGCCAGATTTGCTGTATCTGACCACTGGCGAAGCAATCGCTTCA	2365

RESULT 10

US-10-437-963-77005

Sequence 77005, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 77005

LENGTH: 2780

TYPE: DNA

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT4530\_76946C.1

US-10-437-963-77005

Query Match 26.4%; Score 647.2; DB 19; Length 2780;

Best Local Similarity 59.8%; Pred.No. 5.1e-163;

Matches 1103; Conservative 0; Mismatches 738; Indels 3; Gaps 1;

Qy 125 CACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG 184

Db 457 CACAAGTTCGAGTTCCTCCACCGAAGAAGCTCATCGACGAGTTCACCGAGCCGCGTGAAG 516

Qy 185 GAGACTTTCTTCCTGATGAGCCCTTTGAGGAGGTTCAAGAACCAAGCCAGCTTCCAGAAG 244

Db 517 GAGAGCTTCTTCGCCGACGACCCGTTGCGCAGTACAAAGACCAAGCCGATGTCAAAGAAG 576

Qy 245 TTCATGCTTGGCCTTCAGTCTTCTTCCCATTTTCCGAATGGGCTCCCAATACACCTTT 304

Db 577 GTGTTGATCAGCCTCGAAACTTCTTCCCGGTCTGGAGTGGGGCCGCGCACTACACCTTC 636

Qy 305 CAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCTCTCAG 364

Db 637 CGCAAGTTCAGGGCGACCTCGTCTCGGGCTCACCATTTGCCAGCTCTGCATACCCAG 696

Qy 365 GGCATCAGTTATGCGAAGCTCGCCAACTCCCTCCAAATTTTGGAGCTATATTGAGCTTT 424

Db 697 GACATCGGTTATGCGAAGCTTCGCGCTATTACCAAAATTTATGCACTGTACAGCAGCTTC 756

Qy 425 ATACCAACATTGATTTATGCGATGATGGGTAGCTCAGGGATTTGCGAGTGGGAGCTGTG 484

Db 757 GTACCGCCTTTGATATACGCGATGATGGGAAGTTCCAGGGATATAGCCATTTGTCTCCAGT 816

Qy 485 GCGGTGTGATCGCTTCTGATGGTTTCGATGTTGAGTAATGCGGTTGATCCCAATGAAGAC 544

Db 817 GCGGTGCTTTCGCTGCTGTGCGACTCTCTCTCCAGATGAGTTTGNATCCAGAGAGAT 876

Qy 545 CCAAGAGCTTTACCTCCACTGCTTTTACAGCTTACATTTATTTGCTGGTGTGTTTTTCAGGCT 604

Db 877 CAGGAAGAGTACACCGGCTGCTTCACTGCAACGTTCTTGTGCTGGGTCAACCCAGCA 936

Qy 605 GCTTTGGTCTGTTTAGGTTGGGTTGATCGTGGATTTTCTGTACATGCAACCATATA 664

Db 937 GTGCTCGGATTCCTCAGGCTAGGGTTTATCATAGAGTTCCTTGTGCCATGCTGCTTCGTT 996

Qy 665 GGGTTTCATGGGAGGAGCAGCCAGGTGTGTCTGCGACCACTAAATCGATTTCTTTGGC 724

Db 997 GGAATTCATGGCGGAGCGGCATCAACATTGCCCTTCAGCAGCTTAAAGGCTTCTTTGGA 1056

Qy 725 CTTGAGCATTTTCACCCATGGAGTGAATCATATCATGATGATGCGCTCTGTTTTCACCAA 784



```
Db 1683 ATGG---GTGAAGATGAGGATAACCTGTGCAAGAGCGTCGCGCACGATCTGCAATACTT 1739
Qy 1765 TATAATTGATATAGTGTCTGTTGAAACATTTGATACAAAGTGGAAATAGTATGCTTGAAGA 1824
Db 1740 GGTCTTGTGATCTTGGTGTGTCACCTCTGTCGCAACTCTGGAGTTGGGATGCTACTAGA 1799
Qy 1825 GGTGAAGAAGATTAACAGAGAGAGAGCTACAGCTTGTGTTTGGTCAATCTCTGTAAGTGA 1884
Db 1800 AGTACACAAGAGCCTCGAAGAGAGAGGATCACGATAGCTCTGACGAATCCGAGGCTAGA 1859
Qy 1885 AGTGATGAAGAACTGAACAAATC---GAAGTTCCAAAATCATTTAGGGAAGAAATGGAT 1941
Db 1860 GGTAAACAGAGAGCTGTGCTGTCTGGATAGCTCAGGACATCTTAGGGATGAGTGGGT 1919
Qy 1942 CTATCTGACTGTTTGAAGAGCGGTTGGAGCATGCAACTTCAATCTACGTGCAAGCAAAAC 2001
Db 1920 CTTCTCAGCGTCAAGAGCGCCATCACGGCGTGTGATACGGCGTGCAGATATCCAGAA 1979
Qy 2002 GAACCCAAAGAAAGATGAA 2020
Db 1980 TAAGGGAGAAAGACGAAGTA 1998
```

## RESULT 12

```
US-10-437-963-73410/c
; Sequence 73410, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)/B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 73410
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73694C.1
US-10-437-963-73410
```

```
Query Match 23.3%; Score 571.4; DB 19; Length 2429;
Best Local Similarity 58.1%; Pred. No. 1.3e-142;
Matches 1075; Conservative 0; Mismatches 726; Indels 48; Gaps 2;

Qy 123 TGCACCAAGTGCAGGTTCCACCGCCACAGCGGTTTTCAAGTCTCTAAAGTACTCTTTGA 182
Db 2260 TGCACCAAGTGTGCGTCCGAGAGGAGGTCCGACGGCGAAGCGGCTGAGGCAGCGGTGG 2201
Qy 183 AGGAGCTTTCTCCCTGTATGACCTTTGAGGCGATTCAGACACAGCCAGCTTCCNAGA 242
Db 2200 CGGAGGTGTCTTCTCCCGACGACCGGCTCCACCAAGTTCAAGAACCAAGTCTCGCGCGGC 2141
Qy 243 AGTTTCATGCTTGGGCTTTCAAGTCTTCTTCCCAATTTTCGAATGGGCTCCCAAAATACACCT 302
Db 2140 GGCTGGTGTGCGCGCTGAGTACTTCTTCCCAATCTTCCATTTGGGGCTCCGACTACAGCC 2081
Qy 303 TTCAAGTTCTTGAAGAGTGACCTCATAGCTGGCATFACCATCGCTAGCTTGGCCATTCTTC 362
Db 2080 TCGCCCTCTCCGCTCCGAC----- 2061
Qy 363 AGGCGATCAGTTATGCCAAGCTGGCCAACTCCCTCCCAATTTTGGACTATATTCGAGCT 422
```

```
Db 2060 --GGAATCAGCTACGCCAAGCTCGCCAACTGCCTCCAATCATATTGAGCTATATTTCGAGCT 2003
Qy 423 TTATACCACCATTTATTTATTCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACTG 482
Db 2002 TCGTGGCGCGCTGATCTACTCTGTTGGGTAGCTCGCGGATCTGGCTGTAGGGCGCG 1943
Qy 483 TGGCGGTTGGATCGCTTCTGATGGGTTTCGATGTTGATGTAATCCGTTGATCCCAATGAAG 542
Db 1942 TGTGCGATAGCGTCCGCTGGGTGATGGGGTGCATGCTCCGGCAGGCGGTGTCCCGGACAGG 1883
Qy 543 ACCCAAAGCTTTTACCTCCACCTGGCTTTTACAGCTACATTTATTTGCTGGTGTTTTTCAGG 602
Db 1882 AGCCCATCTCTACCTCCAGCTCGCTTCACTCTCCACTTCTTCGCGGGGCTCTTCCAAG 1823
Qy 603 CTGCTTGGGTCTGTTTATGTTGGGTTGATCGTGGATTTTGTGCATGCAATGCAACATAA 662
Db 1822 CTTCTCCGCTTCTCTCGGCTGGGTTTATCGTGGATTTCTTGTGGAAGCGACGCTGA 1763
Qy 663 TAGGGTTTCAGGAGGAGCAGCACCGTGGTGTCTGCAGCAACTAAATTCGATTTCTTG 722
Db 1762 CGGGATTCATGGCGCGCGCGGCGATCATCGTGTGCTGCAGCAGCTCAAGGGATTTGCTCG 1703
Qy 723 GCCTTGAGCATTTACCCATGGAGCTGATATCATATCAGTGTGCTGTGTTTTCACCC 782
Db 1702 GGATCATCATTTTACGTCGCAGATGGGGTTCGTCCAGGTGATGCACTCGCTCTTCAAGC 1643
Qy 783 AAATCATGATGGAGGTGGGAAAGTGTGTGTGTAGGATGTGTCTTCAATTTTCTTCTCC 842
Db 1642 ACCACGAGTGGGGTGGCAGACCATCTCATGGCGGTGCGCTTCTCGCGGCTCTCC 1583
Qy 843 TTAGCAAGAATACCTTCAGCAAAAACAGCAGGTTTTTTTGGGTGTGAGCAATGGCGC 902
Db 1582 TCACCACAGCCACATCAGCGCCAGGAACCCAAAGCTTTTCTGGGTGTGACGAGCTGCTC 1523
Qy 903 CATTTGAGCTCCGTTATTTGGGAAGTCTCTTGTGTTTATTTTCACTCAGCGCGAGAAGCAGC 962
Db 1522 CACTGACATCAGTGTATCTCTTACCATCATCTCGTTCGTGTCAGCAAGC-----TCATG 1469
Qy 963 GTGTTGAAGTGTAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAATCTGG 1022
Db 1468 GCATCAGTGTATTGGCGATCTCCCAAGGGAATTGAACCCCTCTTCAAGCAATATGCTGA 1409
Qy 1023 TATTTGTGCGCTTACATGACTACAGCTGTCAAACTGGCATTTGTTGGCATCATAT 1082
Db 1408 CTTTCAGTGGCTCTTACGTAGGACTGGCTTAAACACAGGAATTTAGACTGGCATACTAT 1349
Qy 1083 CACTTCGGAAGGAATAGCAGTAGGAAGAGCTTTTGCATGTATAAAATTTACAATATTG 1142
Db 1348 CTCTCACTGAAGGATAGCAGTAGGAGGACATTTGCATCCATCAACACTACAGGTTG 1289
Qy 1143 ATGGCAACAAAGATGATAGCTATTGGGACCAATGAACGCTAGTTGGTTCTTTCACTCTT 1202
Db 1288 ATGGAAACAAGGAGATGATGGCCATTTGGTGTCTATGAACATGGCTGGCTCATGTGCTCCT 1229
Qy 1203 GCTACCTCAACAGAGACCAATTTTCGGTGTGGCTGTGAACTATAAGCTGGATGCAAGA 1262
Db 1228 GCTATGTCAACAGGATCGTTCTCGAGGTGGCGGTGAACCTACAGCGGGGTGCAAGA 1169
Qy 1263 CAGCAGCTTCCAACTATAATTTGTCAGTAAATTTGACATTTGACATTTGTTTCTCGACAC 1322
Db 1168 CGCGGTGTGAACATCTGATGATGGCGTGGCGGTGTGTTGAGCGTGTGCTGTCTGATGC 1109
Qy 1323 CTTTGTTCATTTACACTCCCTGGTGTGCTATCAGCTATTTATCGTATCTGCAATGCTTG 1382
Db 1108 CGCTGTTCCACTACACCCCAACGTCATCTCTCGCGGATCATCATCACCGCGGTGATCG 1049
Qy 1383 GACTCATAGTTATGAAGCAGCCATCCATCTATTTAAGTTTGAATAATTTGACTTTGTGG 1442
Db 1048 GCCTCATGACGTCCCGGGCGCGCAGGCTGTGGAAGGTGCGAAGCTCGACTTCTCTCG 989
Qy 1443 TGTGCAATGATGATACATTTGGCGTGTCTTTGGCAGTGTGTAATTTGGCTTGTAGTCATAG 1502
Db 988 CTTGCAATGGCGCTTCTCTCGGCGTCTCTCTGCTCCGTCAGATGGGCTCGCCATCG 929
```

```
QY 1503 CTATTGTAATATCTGTACTTTCGGGTACTTCTATTATTGCAAGGCCCAAGGACATTCGTTT 1562
Db 928 CCGTCGGCACTCCCTCTTCAAGATCTCTCCAGGTCAACCCGCCCAAAATGGTGTCA 869
QY 1563 TGGCAACATTCANAATTCGTGTATATACCGAAATTTGAGCACTATCAAAATGCAAAAC 1622
Db 868 AGGCGGTGTCCTCCCGGCAACCGGAGCTACCGGAGCATGGCGCAGTACAGGGAGGCCATGC 809
QY 1623 ATGTTCTCGAATGCTAAATCTAGAGATTGATGCACCAATTTACTTTGCCAATGCCAGCT 1682
Db 808 GTGTGCCGTCTTCTCTCGTCCGTCCGGGTCCAGTCCGCCATCTACTTTCGCCAACTCCATGT 749
QY 1683 ATTTAAGAGAAGGATCACAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACAG 1742
Db 748 ACCTCGCGAGAGATCATGAGGTTCTCCGGGAGGAGACGAGCGGCCCAAGTGCA 689
QY 1743 GGGAGACTAGTTTGCAGTATGTTATATATGATATGATGAGTGTGTTGAAACATTTGATACAA 1802
Db 688 ACCAGTCCCTGTGAGATGCATCATCTCGACATGAGTGTGTTGACGATCGACACGA 629
QY 1803 GTGGAATAAGTATGCTTGAAGAGGTGAAGAGATTACAGAGAAAGAGACTACAGCTTG 1862
Db 628 GTGCGCTCGATGCACTTTCAGAGCTGAAGAAGGTGCTGGAAGAAAGAAACATCGAGCTTG 569
QY 1863 TTTTGTCTAATCTGTAAAGTGAAGTGAATGAAGAAACTGAAACAAATCGAAGTTCCAAAATC 1922
Db 568 TGCTGGCCAAACCCGGTTGGATCGGTGACCGAGAGGCTGTACAACTCGGTGTCGGCAGA 509
QY 1923 ATTTAGGAAGAAATGATATCTACTGCTGTTGAAGAGCGGTGGAGC 1971
Db 508 CGTTCCGCTCAGACCGCGTGTCTTTCAGCGTCGCCGAGCGGTCCGGGC 460
```

## RESULT 13

```
US-10-425-114-15719
; Sequence 15719, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaka, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15719
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3060-059-F10_FLI
US-10-425-114-15719
```

```
Query Match 23.2%; Score 567; DB 18; Length 2656;
Best Local Similarity 57.0%; Pred. No. 2.1e-141;
Matches 1060; Conservative 0; Mismatches 795; Indels 6; Gaps 1;
```

```
QY 123 TGCAACCAAGTCGAGGTTCCACGCCACAGCGTTTTTCAAGTCTCTAAAGTACTCTTTGA 182
Db 300 TGCAACAGGTGTCGTCCCGGACGGCGGACGAGCGGAAAGGCGCTGCGGACGGCCTGG 359
QY 183 AGGAGACTTTCTTCCCTGTATGACCTTTTGAGGAGTTCAAGAACAAAGCCAGCTTCCAAGA 242
Db 360 CCAGAGGTGTTCTTCCCGGACGACCCGCTGCACCAAGTTCAAGAACCAAGTCTGTCGGCGGCG 419
QY 243 AGTTCAAGTCTGGCCTTCAGTCTTCTTCCCACTTTTCGATGGGCTCCCAATACACCT 302
```

```
Db 420 GCCTCGTCTGGCGCTGCACACTTCTTCCCACTTCTCCAGTGGGGTCCGCCATACAGCC 479
QY 303 TTCAGTTCTTGAAGCTGACCTCATAGCTGGGATCACCATCGCTAGCTTGGCCATTCCTC 362
Db 480 CGGCGCTCTTGGGCTCCGACCTCGTCCGGGCTTACCAATTCGACGCTCGCCATCCCCG 539
QY 363 AGGCAATCAGTTATGCAAGCTCGCAACCTCCCTCCAAATTCCTTGACTATATTCGAGCT 422
Db 540 AGGGAATCAGCTACGCCAAGCTCGCAACCTCGCCCAATTCGTTGGCTTATATTCAGCT 599
QY 423 TTATACCAACATTTGATATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGAGCTG 482
Db 600 TCGTCCGCGCTCATCTACGCGCTGCTGGGAGCTCGCGGACCTCGCGGTGGGGCCGG 659
QY 483 TGGCGGTTGGATCGCTTCTGATGGGTTGATGTTGAGTAATGCGGTGATCCCAATGAAG 542
Db 660 TGTCCATCCGCTCGCTGGTGGTGCATGCTCCGGGACGCGGTGTCGCGGAGCAGC 719
QY 543 ACCAAAGCTTTTACCTTCCACCTGGCTTTTACAGCTACATTTATTTGCTGGTGTGTTTT 602
Db 720 AGCGCTCTCTACCTGACGCTGGCTTCAACGCCACCTTCTTCGCGGCGTCTTCCAGG 779
QY 603 CTGCTTGGGCTCTGTTTAGGTTGGGTTGATCGTGGATTTCTGTACATGCAACCAATAA 662
Db 780 CGTCCCTGGGATTCCTCAGGCTGGGCTTCATCGTGGACTTCTGTCCAAAGGCGACGCTGA 839
QY 663 TAGGGTTTCATGGGAGGAGCAGCAGCGTGGTGTGCTGAGCAACTTAAATCGATTCTTG 722
Db 840 CGGCTTTCATGGCGGCGCGCGCTCATCGTGTGCTGAGAGCTCAAGGGCTGCTCG 899
QY 723 GCCTTGAGCAATTTCAACCAATGGAGCTGATPATCATATCAGTGATGCGCTCTGTTTT 782
Db 900 GCATCTCCCACTTCACTCCCATGGGATTCTCGAGCTCATGGCTCGTCTCAACC 959
QY 783 AAACCTCAGTGGAGGTGGAAAGTGTGTTGTTAGGATGTGCTTCATTTTCTTCTCCTCC 842
Db 960 GCCACGACGAGTGGAGTGGCAGACATCGTATGGGCTCCGCTTCTTCGCACTCCTCC 1019
QY 843 TTAGCACAAAGTACTTTCAGCAAAAAACGACCAAGGTTTTTTGGGTGTGAGCAATGGCGC 902
Db 1020 TCCTCAGCGCCAAATCAGCGCCAGGACCCAAAGCTTTTCTGGGTATCAGCAGGTGCTC 1079
QY 903 CATTCAGCTCCGTTATTTGGGAAGTCTCTTGTTTATTTTCACTACGCGGAGAAAGCAG 962
Db 1080 CCTCGCGTCGTGTATCTCCACCATCTCTCTTCTTCTTGGAAATCCCCAGCATCA 1139
QY 963 GTGTTGAAGTATAGGAGAACTGAAGAGGTTTGAATCCACATCAGTCAACAATCTGG 1022
Db 1140 GTGTT-----ATTGGCATCCTCCCGAGGGAGTGAACCTCCTTTCGGGAAACATGCTCA 1193
QY 1023 TATTTGTGTCGCTTACATGACTACAGCTGTCAAACTCGCAATGTGCTTGGCATCATAT 1082
Db 1194 GCTTACGGGCTCTTATGTGGCGCTGACGATCAAAACCGGGATCATGACAGGCATCTGT 1253
QY 1083 CACTTTCGGAAGGAATPAGCAGTAGGAAGAGCTTTTGCAATGTATAAAATTTACAATTTG 1142
Db 1254 CTTTAAACGAAGAGGATCGCAGTGGGAGGACCTTCGCGTCCATCAACAACCTACAGGTG 1313
QY 1143 ATGGCAACAAAGATGATAGCTATTTGGACCATGAACGTAGTTGGTTCTTTCACCTCTT 1202
Db 1314 ACGGGAACAAAGAGATGATGGCGATCGGGCTGATGAACATGGCGGGCTCTCGGCTCTCT 1373
QY 1203 GTTACTCTCAACAGGACCATTTTTCGCTTCGGCTGTGAACCTATAACGCTGGATGAAGA 1262
Db 1374 GTTACGTGACAGCGGGTCTTCTCCCGTTCGGCGTGAACCTACAGCGGCGGCTGACGA 1433
QY 1263 CAGCAGCTTCCAAATPATAATGTCACTTTCAGTAAATGTTGACATTTGTTATTTCTTGAC 1322
Db 1434 CGGCGCTGTCCAAACGTGATGGCGGCGGCTGTGGTGAACGCTGTGTTCTTCTCATGC 1493
QY 1323 CTTTGTTCATTTACACTCCCTTGGTGTCTATCAGCTATTATCTATCTGCAATGCTTG 1382
Db 1494 CGCTGTTCCACTACACCCCGGAACGTGATCCTGGCGCGATCATCATCACGGCGGTGG 1553
```

Qy	1383	GACTCATAGATTATGAAGCAGGCATCCATCTATTAAAGTTGACAAAATTGACTTTGTGG	1442
Db	1554	GGCTGGTGACGTCGCGCGCGCCGACGCTGTGGAAGGTGACAAGCTGGACTTCTCTGG	1613
Qy	1443	TGTGCATGAGTGATACATTTGGCGTGGTCTTTGGCAGTGTTCGAAATTGGCTTAGTCATAG	1502
Db	1614	CGTGGTGGCGCGCTTCCTCGCGGTGCTGCTGGTGTGCAGACGGGCTTGGGGCTCG	1673
Qy	1503	CTATTGTAAATCTGTACTTCGGGTGACTTCTATTATTGCAAGGCCAAGGACATTCGTTT	1562
Db	1674	CCGTGGGCATCTCGCTCTTCAAGGTCCTGCTGCAGGTCAACCGCCCAACGTCGTGGTGG	1733
Qy	1563	TGGGCAACATTTCCAAATTCGTGATATACGAAATGTTTGAGCACATCAAAATGCAAAAC	1622
Db	1734	AGGACCTCGTCCGGGACGCAGACTACCGCAGCGTGGCGAGTACCGGAGGCGGTCC	1793
Qy	1623	ATGTTCTTGGAATGCTAATCTTAGAGATTGATGCACCAATTTTACTTTGCGCAATGCCAGCT	1682
Db	1794	GCGTGCCGGGGTTCCTCTCGTCCGGCGTCGAGTCGCGCTCTACTTCGCGCAACTCCATGT	1853
Qy	1683	ATTTAAGAGAAAGGATCAACAGGTGGATTGATGAAGAAGAAGAAAGATTAAAGCTACAG	1742
Db	1854	ACCTGGTGGACGGGTCAATGCTACTCCTCGCGACGAGAGAGCGCGCGCTCAAGTCCA	1913
Qy	1743	GGGAGACTAGTTTGCAGTATGTTTAAATTGATATAGTGTCTGTTGCGAAACATTTGATACAA	1802
Db	1914	ACCACCCCTCCATCCGATCGCTCGTCTCGCATGGGCGCGCTCGCGCGATCGCACAGA	1973
Qy	1803	GTGGAATAAGTATGCTTTGAAGGCTGAAGAGATTTACAGAGAGAGAGGCTACAGCTTG	1862
Db	1974	GGGTCTAGACCGCTGTCCGAGCTCAAGAAAGTCTCTGGAACAAAGAAACATTCGAGCTGG	2033
Qy	1863	TTTTTGGTCAATCCTGTAACTGAAGTGATGAAGAACTGAACAAATTCGAAGTTCCAAATC	1922
Db	2034	TGCTTGGCAACCCCGTGGGGTCCGTTGGCGGAGAGGATGTTCAACTCGCGCGTGGGCGAGA	2093
Qy	1923	ATTTTAGGGAAGAAATGGATCTATCTGACTGTTGAGAGGCGGTTGCGAGCATGCACTTCA	1982
Db	2094	GCCTCGGCTCGGGCGGCTCTTCTTCAGCGTAGCGGAGGCGCTCGCGGGGGCGTGCA	2153
Qy	1983	A	1983
Db	2154	A	2154

## RESULT 14

```

US-10-425-115-75763
; Sequence 75763, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 75763
; LENGTH: 2713
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577 169123C.1
US-10-425-115-75763

```

Query Match 23.2%; Score 567; DB 20; Length 2713;  
Best Local Similarity 57.0%; Pred. No. 2.1e-141;  
Matches 1060; Conservative 0; Mismatches 795; Indels 6; Gaps 1;

Qy	123	TGCAACAAAGTCGAGGTTCACCGCCACAGCCGGTTTTTCAAAGTCTCTAAAGTACTCTTTGA	182
Db	357	TGCACAAGGTGTGCGTCGCCGAGCGCGGACGACGGCGAAGGGCGCTCGCGCAGCGCGCTGG	416
Qy	183	AGGAGACTTTCTCCCTGTATGACCCCTTTGAGGCGAGTTCAAGAACAAGCCAGCTTCCNAGA	242
Db	417	CCGAGGTGTTCTTTCCCGGAGCGACCGCTGCACAGTTTCAAGAACCAAGTCTCTCGCGCGCGC	476
Qy	243	AGTTTCATGCTTGCGCTTCAGTCTCTCTTCCCACATTTTCGAATGGGTCCCAAAATACACCT	302
Db	477	GCCTCGTGTGGCGCTGCATACTTCTTCCCCATCTTCAGTGGGGTCCGCTCACAGCC	536
Qy	303	TTCAAGTTCTTGAAGCTGACCTCATAGCTGGCAATCACCATCGTAGCTTGCCCAATTCCTC	362
Db	537	CGCGCTCCTCGGCTCCGACTCGTCGCGCGCTCACCAATTGCGACGCTCGCCATCCCGC	596
Qy	363	AGGCGATCAGTTATGCCMAAGCTCGCAACCTCCCTCCAAATCTTTGGACATATATTCAGACT	422
Db	597	AGGGAATCAGCTACGCCAAGCTCGCCAACCTCGCGCAATCGTTGGGCTATATTCAGACT	656
Qy	423	TTATACCACCAATGATTTATTCGATGATGGTAGCTCGAGGGATTTGGCAGTGGGGACTG	482
Db	657	TCGTGCGCGCGCTACTACGCGCTGCTGGGAGCTCGCGGACCTTGGCGTGGGGCCGG	716
Qy	483	TGGCGGTTGGATCGCTTCTGATGGGTTGATGTTGAGTAATGCGGTTGATPCCCAATGAAG	542
Db	717	TGTCATCGCGTCTGCTGGTGATGGGTCCATGCTCCGGGACGCGGTGTCGCCGACGAGC	776
Qy	543	ACCCAAAGCTTTAGCTCCACCTGGCTTTGACAGCTACATATTTGCTGTGTGTTTTTCAGG	602
Db	777	AGCGCTCTCTACCTGCACTGGCTTCAACGCCACCTTCTTCGCGGGCGTCTTCCAGG	836
Qy	603	CTGCTTTGGGTCGTTTTAGTGTGGGTTGATCGTGGATTTTGTGCACATGCAACCATAA	662
Db	837	CGTCCCTGGGATTCCTCAGGCTGGGCTTATCGTGGACTTCCTGTCCAAGCCAGCGCTGA	896
Qy	663	TAGGGTTCAATGGGAGGAGCAGCACGGTGGTGTGTCTGCAACAACTAAAATCGAATCTTG	722
Db	897	CGGCTTCATGGCGGCGCGCGCTCATCGTGTGCTGCAAGCAGCTCAAGGGCGCTGCTCG	956
Qy	723	GCCTTGAGCATTTTCACCCATGGAGCTGATATCATATCAGTGATGCGCTCTGTTTTCAACC	782
Db	957	GCACTCTCCCACTTCACCTCCCAATGGGATTCCTCGACGTGATGCGCTCGCTCGTCAACC	1016
Qy	783	AAACTCATGATGGAGTGGGAAAGTCTGTGTTAGGATGTGTCTTCATTTCTTCTCTCC	842
Db	1017	GCACGACGAGTGGAGTGGCAGACGATCGTCAITGGCTCCGCTTCTCGCCATCTCTCC	1076
Qy	843	TTAGCAAGAATACTTTCAGCAAAAAACGACCAAGGTTTTTTGGGTGTCTCAGCAATGGCGC	902
Db	1077	TCCTCAGCGCCAAATCAGCGCCAGGAAACCAAGCTTTTCTGGGTATCAGCAGGTGCTC	1136
Qy	903	CATTGACGTCGGTTATATTGGGAAGTCTCTTGGTTTATTTCTACTCAGCGCGAGAAGCAGC	962
Db	1137	CCCTGGCGTCCGTTGATCATCTCCACCATCCTCTCTTCATCTGGAATATCCCCAGCATCA	1196
Qy	963	GTGTTGAAGTATAGGAGAACTGAAGAAGGGTTTGAATCCACCATCACTCACAAAATCTGG	1022
Db	1197	GTGT-----ATTGGCATCTCCCAAGGGAGTGAACCTCTCTTGGGAAACATGCTCA	1250
Qy	1023	TATTTGTGCGCCTTACATGACTACAGCTGTCAAAACTGGCATTGTCTGTTGGCATCATAT	1082
Db	1251	GCITCAGCGCTCTATGTGGCGCTGACGATCAAAACCGGATCATGACAGGCATCTGT	1310
Qy	1083	CAC TTGGGAAGGAATAGCA GTTAGGAAGAGCTTTTGG AATGTAT AAAAT TACAATTTG	1142
Db	1311	CCTTAA CAGAAGGGATCGCAGTGGGAGGACCTTCGCTC CATCAACA ACTACAGGTGG	1370
Qy	1143	ATGCGACAAGAGATGATAGCTTATTTGGGACCATGAACGTAGTTGTTCTTTTCACTCTT	1202
Db	1371	ACGGGAACAAGGAGATGATGGCGATCGGGCTGATGAACATGGCGGGCTCTCTGCGCTCT	1430
Qy	1203	GCTACCTTCAACAAGGACCAATTTTCGCGTTTCGGCTGTGAACATATAACGCTGGATGCAAGA	1262



```
Db 1431 ||||| GCTACGTGACGCGGGTCTTCTCCCGTGGCGGTGAACATACAGCGGGGTGCAGGA 1490
Qy 1263 CACGAGCTCCCAACATATATACATCTGCTGACGTAATGTTGACATGTTATCTTCACAC 1322
Db 1491 CGGCGGTGTCACACGTCGTGATGGCGGCGGTGCTGTGGTGCAGCTGCTGTTCCCTATGC 1550
Qy 1323 CCTTGTTCATTAACATCCCTCGTGTGCTATCAGCTATTATCGTATCTGCAATGCTTG 1382
Db 1551 CGCTGTTCCACTACACCCGAACTGATCTGCGGGGATCATCATCAGCGGCGTGGTGG 1610
Qy 1383 GACTCATAGATTATGAAGACGCCATCCATCTATTAAAGTTGACAAATTTGACTTTGTGG 1442
Db 1611 GGTGTGTGACGTCGCGCGGCGCCAGCTGTGGAAGTGGACAAAGCTGGACTTCCTCG 1670
Qy 1443 TGTGCATGAGTGACATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGCTTAGTCATAG 1502
Db 1671 CGTGCCTGGCGGCTTCTCGGCGTGTGCTGTGTGTCGTCGACAGCGGCGCTGGGCGTGG 1730
Qy 1503 CTATTGTAATATCTGTACTTTCGGGTACTTCTATTATTGCAAGGCCAAAGACATTCGTTT 1562
Db 1731 CGCTCGGCATCTCGCTCTTCAAGTCTCTGTCAGGTCAACCGCCCAACGTCGTGGTGG 1790
Qy 1563 TGGGCAACATTCGTAATCTGTGATATACGAAATGTTGAGCACTATCAAAATGCAAAAC 1622
Db 1791 AGGCGCTCGTCCCGGGGACGACAGACTACCGCAGCTGCGCAGTAGTACCAGCGAGCGCTCC 1850
Qy 1623 ATGTTCTCGAATGCTAATCTAGAGATTGATGCACCAATTTACTTTGCCAATGCCAGCT 1682
Db 1851 GCGTCCGGGCTTCTCGTGTGCGGCTGAGTCCGCGCTACTTTCGCCAATCCATGT 1910
Qy 1683 ATTTAAGAGAAGGATCACAAAGTGGATTGATGAAGAAGAAAGAAATTAAGCTACAG 1742
Db 1911 ACCTGTGGAGCGGTCATCGCTACCTCGCGACGAGGAGGAGCGCGCTCAAGTCCA 1970
Qy 1743 GGAGACTAGTTTGCAGTAGTTATTAATTAATGATATGATGCTGTTGGAACATTTGATACAA 1802
Db 1971 ACCACCCCTCCATCCGATGCTGCTCTCGACATGGCGCGCTGCGCGGATGTCACACGA 2030
Qy 1803 GTGGAATAGTATGCTTGAAGAGTGAAGAGTACAGAGAGAGAGAGAGTACAGCTTG 1862
Db 2031 CGCGTCTAGACGCGTGTGCGAGCTCAAGAAGTCTCGGACAAAGAAACATCGAGCTGG 2090
Qy 1863 TTTTGTCTAATCTGTAAGTGAAGTGAAGAAACTGAACAAATCGAAGTTCGAAATC 1922
Db 2091 TGCTTGCCAAACCGGTGGGTGCTGTCGCGAGAGGATGTTCACTCGGCGGTGGCGGAGA 2150
Qy 1923 ATTTAGGGAAGAAATGATCTACTGCTGTTGAAGGCGGTTGGAGCAATGCAACTTCA 1982
Db 2151 GCTTCGGGTGGGCGCGCTTCTTTCAGCGTAGCGGAGGCGGTGCGCGGGGGCGGTGCA 2210
Qy 1983 A 1983
Db 2211 A 2211
```

RESULT 15  
US-10-424-599-32308  
; Sequence 32308, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 32308  
; LENGTH: 536

```
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129179C.1  
US-10-424-599-32308  
  
Query Match 21.2%; Score 520.2; DB 18; Length 536;  
Best Local Similarity 98.5%; Pred. No. 3.3e-129;  
Matches 525; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
Qy 7 GCTAGCTCGCACATTAAGTTATATAACACATATTTCTTCTGCTAGAAATACTATTATTCA 66  
Db 4 GCTAGCTCGCACATTAAGTTATATAACACATATTTCTTCTGCTAGAAATACTATTATTCA 63  
  
Qy 67 AGATATGGGAGTGTAGATTATAGTACCCCTTTGGGCATGAACAACTTTGAGAGAGTGA 126  
Db 64 AGATATGGGAGTGTAGATTATAGTACCCCTTTGGGCATGAACAACTTTGAGAGAGTGA 123  
  
Qy 127 CCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAAAGTCTCTAAAAGTACTCTTTGAAGGA 186  
Db 124 CCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAAAGTCTCTAAAAGTACTCTTTGAAGGA 183  
  
Qy 187 GACTTTCTTCCCTGATGACCCCTTTGAGGAGTTCAGAACAGCCAGCTTCCAAAGAGTT 246  
Db 184 GACTTTCTTCCCTGATGACCCCTTTGAGGAGTTCAGAACAGCCAGCTTCCAAAGAGTT 243  
  
Qy 247 CATGCTTGGCCCTTCAGTTCTTCTCCCAATTTTGAATGGGCTCCCAAAATACACCTTTCA 306  
Db 244 CATGCTTGGCCCTTCAGTTCTTCTCCCAATTTTGAATGGGCTCCCAAAATACACCTTTCA 303  
  
Qy 307 GTTCTTGAAGCTGACCTCATAGCTGGCATCAACCATCGCTAGCTTGGCCATTCCTCAGGG 366  
Db 304 GTTCTTGAAGCTGACCTCATAGCTGGCATCAACCATCGCTAGCTTGGCCATTCCTCAGGG 363  
  
Qy 367 CATGAGTTATGCCAAGCTGCCAACTCCCTCCCAATTTCTTGGACTATATTTCAGAGCTTTAT 426  
Db 364 CATGAGTTATGCCAAGCTGCCAACTCCCTCCCAATTTCTTGGACTATATTTCAGAGCTTTAT 423  
  
Qy 427 ACCACCATTTGATTTATGCGATGATGGTAGCTCGAGGGATTTGGCAGTGGGACCTGTGGC 486  
Db 424 ACCACCATTTGATTTATGCGATGATGGTAGCTCGAGGGATTTGGCAGTGGGACCTGTGGC 483  
  
Qy 487 GGTTCGATCGCTTCTGATGGGTTTCGATGTTGAGTAATGCCGTTGATGCCAATG 539  
Db 484 GGTTCGATCGCTTCTGATGGGTTTCGATGTTGAGTAATGCCGTTGATGCCAATG 536  
  
Search completed: August 31, 2005, 09:39:23  
Job time : 2775 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 02:28:41 ; Search time 7917 Seconds  
(without alignments)  
11774.586 Million cell updates/sec

Title: US-10-762-049-17  
Perfect score: 2449  
Sequence: 1 gcacgagctagctgcacat.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_ges1.\*  
9: gb\_ges2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	938.4	38.3	2158	3	CNS0A7F1	BX822461 Arabidops
2	670.4	27.4	1959	9	CL961102	CL961102 OsIFCC005
3	655.6	26.8	854	6	CB892638	CB892638 EST645430
4	653.8	26.7	2279	3	AV105934	AV105934 Zea mays
5	605.4	24.7	2001	9	CL959635	CL959635 OsIFCC003
6	581.4	23.7	583	2	AW598447	AW598447 SJ91904.Y
7	570.8	23.3	1812	9	CL969864	CL969864 OsIFCC019
8	567	23.2	567	4	BM093232	BM093232 saJ06610.
9	520.4	21.2	522	4	BG406552	BG406552 sac33f11.
10	516.4	21.1	594	7	CO982632	CO982632 GM89016A2
11	512.6	20.9	2221	3	CNS09YVW	BX832405 Arabidops
12	501.4	20.5	828	7	CV130479	CV130479 X9P02a03
13	493.2	20.1	498	4	B1972185	B1972185 sg988C07.
14	476	19.4	982	7	CK269206	CK269206 EST715284
15	468.2	19.1	583	5	BU090662	BU090662 su07b07.Y
16	461.4	18.8	489	2	BE804764	BE804764 ss45d02.Y
17	459.6	18.8	2321	3	CNS0A1LZ	BX832587 Arabidops
18	448.4	18.3	831	7	CO107436	CO107436 GR_EB003
19	448	18.3	459	5	BU544910	BU544910 GM880003B
20	437.2	17.9	768	7	CN190878	CN190878 UCRCS06.0
21	433.8	17.7	816	6	CB624151	CB624151 OSTIEa12G
22	423	17.3	706	7	CO102483	CO102483 GR_EB002
23	416.6	17.0	738	5	BU025764	BU025764 QHG11F06.
24	411.8	16.8	728	5	BU025129	BU025129 QHG7N01.Y

## RESULT 1

CNS0A7F1

LOCUS

DEFINITION

Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB32C02 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

TITLE

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

JOURNAL

REFERENCE

Unpublished

2 (bases 1 to 2158)

Genoscope.

Direct Submission

TITLE

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis Genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full

length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

Location/Qualifiers

1. .2158

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/strain="Col-0"

/db\_xref="taxon:3702"

FEATURES

source

Location/Qualifiers

1. .2158

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/strain="Col-0"

/db\_xref="taxon:3702"

## ALIGNMENTS

CNS0A7F1 2158 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFB32C02 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).

BX822461 GI:42465936

HTC: GSLT cDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 2158)

Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

Unpublished

2 (bases 1 to 2158)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis Genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full

length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

Location/Qualifiers

1. .2158

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/strain="Col-0"

/db\_xref="taxon:3702"

FEATURES

source

Location/Qualifiers

1. .2158

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/strain="Col-0"

/db\_xref="taxon:3702"

FEATURES

source

Location/Qualifiers

1. .2158

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/strain="Col-0"

/db\_xref="taxon:3702"

		/clone="GSLTFB3ZC02" /tissue type="Flowers and buds" /plasmid="pCMVSPORT 6" complement (1..2158) /gene="At3g51895"			
		Query Match Best Local Similarity 38.3%; Score 938.4; DB 3; Length 2158; Matches 1328; Conservative 0; Mismatches 551; Indels 9; Gaps 5;			
Qy	125	CACCAAGTCGAGGTTCCACCGGCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG	184		
Db	119	CACAAGTGGAGGCTCCACAACCTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAG	178		
Qy	185	GAGACTTTCTCCCTGATGACCCCTTTGAGGACGTTTCAAGAACAGCCGTTCCAGAGAG	244		
Db	179	GAAACTCTGTTTCCAGACGACCTTTTAGACAAATTTAAGAACCAAAATGATCAAGAAA	238		
Qy	245	TTCATGCTTGGCCTTCAGTTCCTTCTCCCAATTTTGAATGGGCTCCCAATACACCTTT	304		
Db	239	TTTGTGTTAGGCTTCAATACTTCTCCGATTTTCGAATGGGCAACAAGTACATCTC	298		
Qy	305	CAGTCTTTGAAGCTGACCTCATAGTGGCATACCATCGCTAGCTTGGCCATTTCTCTAG	364		
Db	299	AAGTCTTTCAAAATCAGATCTCATCGCGGAATCACCATCGTACTAGTCTCGCCATCCCTCAG	358		
Qy	365	GGCATCAGTATGCCAGCTCGCCCACTCCCTCCAAATCTTGGACTATATTGGAGCTTT	424		
Db	359	GGCATCAGTATGCCCAACTTGTAACTTGGCCCCCAATCTTGGCCCTTTATTGAGTTTT	418		
Qy	425	ATACCACCATTTGATGCGATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGCACTGTG	484		
Db	419	GTACCGCAATTTGTATACGCGTGTCTAGGGAGTTCAAGGACTTAGGGGTGGAAACGTT	478		
Qy	485	GCAGTTGGATCGCTTCGTATGGGTTTCAAGTTGAGTAATGCCCTTGAATCCCAATGAAGAC	544		
Db	479	GCAGTTGGCTCTCTGTGTGACAGGTGCAATGCTGAGCAAAAGATGTGATGCTGAGAAAGAT	538		
Qy	545	CCAAAGCTTTACC-TCCACTGGCTTTTCAAGCTGATATATTTTGTGCTGGTGTCTTTCAGGC	603		
Db	539	CCTAAGCTTTTACCATTCACTTCTTTCACCGCCACTTTTTCGCGGGGCTTCTCGAAGC	598		
Qy	604	TGCTCT-TGGCTCTGTTTAGGTTGGGGTTGATCGTGGATTTTCTGTGATGATTTCTGTGATCAACATAA	662		
Db	599	CTCATCTTGGAAATTTTCAGGTTAGGGTTTATAGTGGACTTCTATCGCATGCAACGATAG	658		
Qy	663	TAGGTTTCATGGGAGGACGCCACGGTGTGTCTGACGAACTAAATTCGATTTCTTG	722		
Db	659	TAGGATTCATGGGAGGAGCAGCAACGGTGTGTGAGTCTGCAACAGATTTAAGGGTATTTTAG	718		
Qy	723	GCCTTGACATTTTCAACCATGGAGCTGATATCATATCATGATGCTGCTGTTTTCACCC	782		
Db	719	GAATTCAAATTTTTCACAGACTCAACAAATGTTATCAGTCACTGATCAATTTTCAAA	778		
Qy	783	AAACTCATAGTGGAGTGGGAAGTGTGTGTAGGATGTGCTTCAATTTCTCTCTCC	842		
Db	779	AAACAACAAGTGGAGATGGGAAGTGGCGTTTATGATGTGTTTTCAAATTCATTCATC	838		
Qy	843	TTAGCAACAAGATCTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTGACGAATGGCGC	902		
Db	839	ACACAACAAGATATTTTCAAGCATCAAGCAACCAAAATTTCTTTGGTGGCGCGATGGCTC	898		
Qy	903	CATTGAGCTCCGTATATTGGGAAGTCTCTGTGTTTATTTCATCTCAGCGCGAGAACGACG	962		
Db	899	AAATTGACCAAGTGATTTCTGGCAGTCAATGTGTTTTCATCAACAACAAGAGACATG	958		
Qy	963	GTGTTGAAGTGTATAGGAGAACTGAAGAGGGTTTGAATCCACCATCACTCACAAATCTGG	1022		
Db	959	GTGTTCAAGTTCATAGGGGCCAGAGACAGG---TTGCTTCACTCAACAAGTTCTGATCTCA	1015		
Qy	1023	TATTTGTGCTCGCCTTACATGACTACAGCTGTCAAACTGGCATTTGCTTGGCATCATAT	1082		

1016	TCTTTTCTTTCACATT-CATGTCCACAGCTGTCCACACCTGGGCCACATCAGCTGGGATCATTTG	1073
1083	CACTTGCGGAAGGAATAGCAGTAGGAAGAAGCTTTTGCATATGTATATAAAATTTACAATATTG	1142
1075	CTCTCGCTGAAGAGTAGCAGTGGGAGGAGTTTTCGATGTTCAGAGAACTACAACATAG	1134
1143	ATGGCAACAAAGAGATGATAGCTATTGGGACCATGAAGCGTAGTTGGTTCTTTTACCTCTTT	1202
1135	ACGGGAACAAAGAGATGATAGCTTTTGGAAATGATGAACATCATCGTTGGTTCTTTCACATCTT	1194
1203	GCTACCTCACAAAGGACCAATTTTCGGTTCGGCTGTGAACTATAACGCTGGATGCAAGA	1262
1195	GTACCTCACAAACCGGACCAATTTTCAAGGTTCGGCAGTGAACACTACAACGCGGTTGCAAGA	1254
1263	CAGCAGCTTCCAAACATATATAATGTCACTTGCAGTAATGTGTGACATTTGTATTTCTCTGACAC	1322
1255	CGCAATGTCCAAACATAGTATGGGATTTGGGTATGTTCACACTCTCTTCTCTCACAC	1314
1323	CTTTGTTCCATTACCTCCCTGGTGGTCTATCAGCTATTATTCGTAATCTGCAATGCTTG	1382
1315	CGCTTTTTCATACACACCACTCGTCTCTCTCTGCGCATCATATATCGCAATGCTCG	1374
1383	GACTCATAGATATTGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTTGTGG	1442
1375	GACTCATTTGACTATCAAGCTGCCATCCATCTCTGGAAGTTGCAAGTTTGACTTTCTCTCG	1434
1443	TGTGCATCAGTGCATACATTTGGCGTGGTCTTTTGGCAGTGTGTAAATTTGGCTTTAGTCTAG	1502
1435	TCTGCATGAGCGCCTACGTTGGGTGCTATTTCCGCGAGTTAGAGATTTGGACTGTCGTAG	1494
1503	CTATTGTAATATCTGTACTTCCGGGTACTTCTATTATTGCAAGGCCAAGGACATTCGTTT	1562
1495	CGGTGGCGATATCTATAGCGAGGTTGTCTGTTGTGTGTCAGGCCAAAAAATCGCGGTGA	1554
1563	TGGGCAACATTTCCAAAATCTCTGTGATATACGAAATGTTGAGCACTATCAAAATGCAAAAC	1622
1555	AGGGAAACATACCAAAACAGCATGATCTATAGGAACACTGAGCAGTACCCATCATCAAGAA	1614
1623	ATGTTCTGGAAGTCTAATCTCAGAGATTGATGCAACCAATTTTACTTTGCCAATGCCAGCT	1682
1615	CGTTCTGTTTCTCTATCTTGGAGATTGATGCTCCCATCTACTTTGCTTAATGCGAGTT	1674
1683	ATTTAAGAAAGAGATCAAGGTGGATTGATGAAGAAGAAGAAGAAATTTAAAGCTACAG	1742
1675	ACTTGGTGAAGAATCATAGGTGGATTGATGAAGAAGAAGAAGAGAGAGTTTAAACAATCAG	1734
1743	GGGAGCTAGTTTTCAGCTATGTTATATTTGATATGAGTGTCTGTTGGAAAAATTTGATACAA	1802
1735	GAGAGAGCAGCTTACAAATATATATCTCGATATGTGAGCTGTTGGTAATATCGACACAA	1794
1803	GTGGAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTTG	1862
1795	GCGTATTAGCATGATGTTGGAATTTAAGAAAGTCAATTGACAGGAGAGCGTTAAAGTTGG	1854
1863	TTTTGGTCAATCTGTAAAGTGAAGTGAAGAACTGAAACAAATCGAAGT---TCCAAA	1919
1855	TATTGTCAATCCAAAAGGAGAGGTCTGTAAGAAATTTAAACAGATCCAAATTCATCGGTG	1914
1920	ATCATTTAGGGAAGAAATGATCTATCTGACTGTTTGAAGAGCGGCTTGGAGCATGCAACT	1979
1915	ATCATTTGGGCAAGAGGTGGATGTTCTTAAACGTTAGGAGAAGCAGTGGAGGCTTGTAGCT	1974
1980	TCAATCTACGTGCAAGCAAAACGAACCC	2007
1975	ACATGCTTTCACACGCTTTTAAACCGAACC	2002
RESULT 2		
CL9611102		
LOCUS		
DEFINITION		
OsIFCC005483 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.		
ACCESSION		
CL9611102		
CL9611102		
GSS 21-SEP-2004		

RESULT 2  
CL961102  
LOCUS  
DEFINITION  
ACCESSION

CL961102 1959 bp DNA linear GSS 21-SEP-2004  
Oe18CC005483 Oryza sativa Express Library Oryza sativa (indica  
cultivar-group) genomic, genomic survey sequence.  
CL961102

VERSION	CL9611102.1	GI:52376926	
KEYWORDS	GSS.		
SOURCE	Oryza sativa (indica cultivar-group)		
ORGANISM	Oryza sativa (indica cultivar-group)		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.		
REFERENCE	1 (bases 1 to 1959)		
AUTHORS	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. W. and Wang, J.		
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped.		
FEATURES	Location/Qualifiers		
source	1..1959 /organism="Oryza sativa (indica cultivar-group)" /mol_type="genomic DNA" /db_xref="taxon:39946" /clone_lib="Oryza sativa Express Library" /note="Oryza sativa exon trapped genomic sequences"		
ORIGIN			
Query Match	27.4%; Score 670.4; DB 9; Length 1959;		
Best Local Similarity	60.3%; Pred. No. 7.2e-164;		
Matches 1109; Conservative	0; Mismatches 731; Indels 0; Gaps 0;		
Qy	126	ACCAAGTCGAGGTCCACCCGACGCGTCTTTCAGTCTCTAAAGTACTCTTTGAGG	185
Db	83	ACATGTCGGCGCGCCGCCGAAGAAGAACCTCTCGCGGAGTTCGCCGACCGTGAAG	142
Qy	186	AGACTTTCTTCCCTGTAGTACCTTTGAGGCGAGTTCAGAGAACAGCCAGCTTCAAGAAGT	245
Db	143	AGAGTTCTTCTCGAGAGCGGATGCGCGGTACAGGACCGAGGTCAGGAGC	202
Qy	246	TCAATGTCGCTTCAGTCTTCTTCCCTCCATTTTCGAATGGGCTCCCAATATACACCTTC	305
Db	203	TGTGGCTCGCCCTGCAGCAGTCTTCCCGGTGTTCGAATGGGCGAGCAATACACCTCG	262
Qy	306	AGTTCCTTGAAGCTGACCTCATAGCTGGCATCCATCGCTAGCTTGGCCATTCCTCAGG	365
Db	263	CCAAGTTCAAGGGCGACCTCAITTCGCGGCTCACCTTTCGACGCTCGTCATACCTCAG	322
Qy	366	GCATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTGGACTATATTCGAGCTTTA	425
Db	323	ACATCGCTACGGAAGCTTGCTAACCTTGCCACAGAGATTTGGCTGCACAGTAGCTTCG	382
Qy	426	TACACCAATTGATTTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGAGCTGTGG	485
Db	383	TCCGCCATTGATATACGCTCTGATGGGCACCTCAAGGAGCTAGCAATGGGTCCAGTGG	442
Qy	486	CGGTGGATCGCTTCTGATGGGTTCGATGTTGAGTATGCGGTGATCCCAATGAAGACC	545
Db	443	CCGTCATCTCACTGCTCTGGTACTCTCTCCAGGAGGAGATGACTCAAGAAGAAC	502
Qy	546	CAAGCTTTTAACTCCACTTTCACAGCTACATTTATTTGCTGGTGTGTTTTCAGGCTG	605
Db	503	CGCTAGATTACAGAGGCTCGCTTTCACAGGACCTTCTTTGACGAGTGCACAGCGCG	562
Qy	606	CTTGGGTCTGTTTATGATGGGTGATTCGTTGATTTCTCTCATATGCAACCATATATAG	665
Db	563	CGCTTGGTTCTGCGAGCTAGGGTTTCATAGGTTCTCTGCTCATGCTGCCATCATCG	622
Qy	666	GGTTCATGGGAGNGCAGCCACGGTGGTGTGTCGACGAACTAAATTCGATCTTGGCC	725
Db	623	GATTCATGGCGCGCGCCCATCACCATTCTTCAGCAGCTTAAAGGCTTCTTGGA	682
Qy	726	TTGAGCATTTTACCCTTGGAGCTGATATCATATCAGTGATGCGCTCTGTTTTCACCCAA	785
Db	683	TTGCAAAATTTACCAAGAGACTGACATCATCTCGGTATGAAATCAGTCTGGGAATG	742
Qy	786	CTCATGATGAGGTGGGAAGTGTGTGTAGATGTGTCTTCATTTTCTCTCTCTTA	845
Db	743	TTCCACCATGGGGCAATGGAACTGCGCAGACAATATTGATCGGAGCATCATTTTGGCATT	802
Qy	846	GCAACAAGATACTTTCAGCAAAAACGCAAGGTTTTTTTGGTGTGTCAGCAATGCGGCAT	905
Db	803	CTCTGTTGTTCCCAAGGCAAGAGAACAAAGAGCTCTTCTGGTCCCAGCAATTTGCACAC	862
Qy	906	TGACGTCCTGTTATATTGGGAAGTCTCTTGGTGTATTTTCACTCACCGCAGAGCACGGTG	965
Db	863	TCATTTTGGTGATCAATTTCAACTTTGTTGCTCTACATCACTCTGCTGACAAAACAGGCG	922
Qy	966	TTGAAGTGATAGGAGAACTGAAAGAGGTTTGAATCCACCATCACTCACAATCTGGTAT	1025
Db	923	TGCAATTTGTAAAAAAGCTCAAGAAAGGCATCAATCCACCTTCAGCTAGCTGATATTT	982
Qy	1026	TTGTGTCGCTTACATGACTACAGCTGTCAAACTGTCATTTGCTGTTGGCATCATATCAC	1085
Db	983	TCACTGGCCCATCTTCTCAAAGGATTCAAAATTTGAGTAGTAGTGGAAATGATAAGCC	1042
Qy	1086	TTGCGGAAGGAATAGCAGTAGGAAGAGCTTTTCAATGTATATAAAATTTACAATATTGATG	1145
Db	1043	TTACGGAAGCGATTGCGAGTTGGAAGAACAATTTCTGGATTGAACGATTACAGATAGATG	1102
Qy	1146	GCAACAAAGAGATGATAGCTATTGGGACCAATGAACGTAGTTGGTCTTTCACCTCTTGCT	1205
Db	1103	GGAACAAAGAAATTTGGCTCTAGGAACCAATGAATGTGGTCTCAATACGCTTTGCT	1162
Qy	1206	ACCTCAACAGGACCAATTTTGGCGTTTCGCTGTAACATAACGCTGGAATGCAAGACAG	1265
Db	1163	ATATAGCCACAGTGGTTTTCAGCATCAGAGTCAATTCGATGGCTGGAGGTAAACAC	1222
Qy	1266	CAGCTTCCAACTTATTAATGTCACTTCGAGTAATTTGACATTTGTTATTCCTGACACCT	1325
Db	1223	CAATGTCCTCAATTTGTTATGTCACTGTAGTATTTGCTGCACTCTCTGTGATCACTCAT	1282
Qy	1326	TGTTCAATTACATCCCTCGTGGTGTCTATCAGCTATTTATCGTATCTGCAATGCTGGAC	1385
Db	1283	TGTTCAAGTACACGCGCAATGCCCAATTTCTTCATCATCATATCAGCAGTGTGGCC	1342
Qy	1386	TCATAGATTATGAAGCAGCCATCCATCTATTTAAAGTTGACAAATTTGACTTTGTGTGT	1445
Db	1343	TATTTGACTTTGAAATCAGCCCTACCTTATCTGGAAGTTGATAAGTTGGACTTTCATGGCT	1402
Qy	1446	GCAATGATGATACATTTGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAATAGCTA	1505
Db	1403	GCTTGGGGCATTCCTTTGGAGTAAATTTTTCATCTGAGGATATGCTGCTCATTCGCG	1462
Qy	1506	TTGTAATATCTGTACTCTCGGTACTCTTATTTATGCAAGGCCAACAGACATTCGTTTGG	1565
Db	1463	TTGTAATATCCTTAATCAAAAGTTCTGCTCCATGTAACAGCCCAAGGACGCTTTACTTG	1522
Qy	1566	GCAACATTTCCAAATCTGTGATATACGAAATGTTGAGCACTATCAAAATGCAAAACATG	1625
Db	1523	GCAACCTTCCAGAACGATTTCTATAGAAATTTGAAACAAATATCCAGAACTACCAAG	1582
Qy	1626	TTCTCTGGAATGCTAAATCTAGAGATTGATGCACAATTTTACTTTTGCAATGCCAGCTATT	1685
Db	1583	TGCCAGGATGCTAAATTTGTAAGAGTGGACTCAGCAATATACTTCCAAACTCCAAATTATG	1642
Qy	1686	TAAGAAAGGATCAACAGGTGATGTAAGAGAGAAAGAAATTAAGAGCTACAGGG	1745
Db	1643	TTAAAGAAAGAAATGCTCAGATGCTGAGAGATGAGGAAGAACATCAAAAGGAAACAGAGT	1702
Qy	1746	AGACTAGTTGCAAGTATGTTATTAATTTGATATGATGCTGTTGGAACATTTGATACAGTG	1805

```
Db 1703 TACCAAAAATTGAGTTTCTGATTGTTGACCTATCTCTCTGTAATGATATTGACACAAGTG 1762
Qy 1806 GAATAAGTATGCTTGAAGAGGTGAAGAGATTACAGAGAGAGAGAGCTACAGCTTGTTT 1865
Db 1763 GAATCCATGCTTCAAGAGATTGTTGAGGACATTGAAAGGCCAGATTCACTGATTT 1822
Qy 1866 TGGTCAATCTCTGTAAGTGAAGTGAAGAGAACTGAACAAATCGAAGTTTCCAAAATCAT 1925
Db 1823 TGCCCAATCTGGGGCGGTGATCCAAAAGCTCCGGTCAGCAAAATTCACAGAGCTCA 1882
Qy 1926 TAGGGAAGAAATGATCTATCTGACTGTTGCAAGAGGCCGT 1965
Db 1883 TTGGTGAAGAAAAGATATGCTGACACAGTTGGTGACGCCGT 1922

RESULT 3
LOCUS CB892638 854 bp mRNA linear EST 24-APR-2003
DEFINITION EST645430 HOGA Medicago truncatula cDNA clone HOGA-19G10, mRNA
sequence.
ACCESSION CB892638
VERSION CB892638.1 GI:30099806
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 854)
AUTHORS Hahn,M.G., Ojanen-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,
Utterback,T., Cho,J. and Fraser,C.M.
TITLE ESTs from roots of Medicago truncatula treated with
oligogalacturonides of DP 6-20
JOURNAL Unpublished (2001)
COMMENT Contact: Michael G. Hahn
Complex Carbohydrate Research Center
University of Georgia
220 Riverbend Road, Athens, GA 30602-4712, USA
Tel: 706-542-4457
Fax: 706-542-4412
Email: hahn@ccrc.uga.edu
TIGR sequence name: MTMCV417K
More information is available at: www.medicago.org
Seq primer: SKnod (CTA GAA CTA gtg gat CC).

FEATURES
source
1..854
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="HOGA-19G10"
/tissue_type="3 day old seedling roots"
/dev_stage="24 hours after treatment in the dark at 26 C
with 0.5 mg/ml oligogalacturonides (DP 6-20) in the
presence of 100 ug/ml Gentamicin"
/lab_host="XL0LR"
/clone_lib="HOGA"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in SOLR cells."

ORIGIN
Query Match 26.8%; Score 655.6; DB 6; Length 854;
Best Local Similarity 85.5%; Pred. No. 4.2e-160;
Matches 730; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
Qy 273 CGATTTTCGATGGGCTCCAAATACACCTTTCAGTTCTTGAAGCTGACCTCATAGCTG 332
|||||
```

```
Db 1 CCATATTTGAATGGGCACCAAGCTACATTTTTCACCTTTTGAAATCTGATCTCATAGCTG 60
Qy 333 GCATCACCATCGCTAGCTTGGCCATTTCCTCAGGGCATCAGTTTATGCAAGCTCGCCAAACC 392
Db 61 GCATCACTATTGCTAGCTTGGCTATTTCCTCAAGGCATTAGTTATGCCAAACTTGGCAACC 120
Qy 393 TCCCTCCAATTTCTGGACTATATTGAGCTTTATACCACTTGAATTTATGCGATGATGG 452
Db 121 TCCCTCCAATTTCTGGACTATATTCAAGCTTTATACCACTTGAATTTATGCAATGATGG 180
Qy 453 GTAGCTCGAGGAGTTTGGCAGTGGGACATGTGGCGGTGGATCGCTTCTGATGGGTTCA 512
Db 181 GTAGTTCAAGGAGTTTGGCAGTTGGAACTGTGGCAGTTGGATCTCTCTCTGATGGGTTCA 240
Qy 513 TGTGTAGTAATGCCGTTGATCCCAATGAAGACCCAAAAGCTTTACCTCCACCTGGCTTTCA 572
Db 241 TGTGTGCCAATGAGGTTAATCTTACCCAAAATCCAAAAGCTTTTCTGCACTTGTCTTTCA 300
Qy 573 CAGCTACATTTTGTGTGTTGTTTTCAGGCTGCTTGGGTCCTGTTAGTTTGGGTTGA 632
Db 301 CAGCTACATTTTGTGTGTTGTTTTCAGGCTGCTTGGGTCCTTCTTAGTTAGGTTTA 360
Qy 633 TCGTGATTTTCTGTCATGCAACATTAATAGGTTTCATGGGAGAGCAGCACCGTGG 692
Db 361 TCGTGATTTTCTGTCATGCAACATTAATAGGTTTCATGGGAGAGCAGCACCGTGG 420
Qy 693 TGTGTCTGCAGCAACTAAATCGATTCTTGGCTTGGAGCATTTTCACTGAGCTGATA 752
Db 421 TTGTGCTACAGCAGCTGAAATCAATTTAGGACTTTGAACATTTTCACTGCTGATA 480
Qy 753 TCATATCAGTGTGCTGCTGTTTTCACCCAACTCATGAGTGGAGTGGGAAAGTCTG 812
Db 481 TTGTATCGTGTGCTGCTGTTTTCACCCAACTCATGAGTGGAGTGGGAAAGTCTG 540
Qy 813 TGTAGGATGCTCTTCAATTTCTCTCTTAGCACAAGATACCTCAGCAAAAACGAC 872
Db 541 TGTGGGATTTTGTCTTCATTTTCTCTCTTGTACAAAGATACCTCAGCAAAAACGAC 600
Qy 873 CAAGGTTTTTTGGGTGTCAGCAATGCGGCCCATTCAGCTCCGTTATATTGGGAAAGTCTCT 932
Db 601 CCAAGTTCTTTGGGTGTCAGCAATGCGGCCCATTCAGCTCCGTTATATTGGGAAAGTCTCT 660
Qy 933 TCGTTATTTTCACTCAAGCCGAGAGACAGGTTGTTGAAGTATAGAGAACTGAAGAAGG 992
Db 661 TCGTTATTTTCACTCAAGCCGAGAGACAGGTTGTTGAAGTATAGAGAACTGAAGAAGG 720
Qy 993 GTTTGAATCCACCATCACTCAAAATCTGGTATTTGTGTCGGCTTACATGACTACAGCTG 1052
Db 721 GACTAAATCCACCATCACTCAAGATTTTGGTATTTGTGTGTCAGCATATATGACACAGCTA 780
Qy 1053 TCAAAACTGGCATTTGCTTGGCATCATATCACTTGGCAAGGAATAGCAGTAGGAAGAA 1112
Db 781 TCAAAACAGGCTTATTTGTCATTTAGCTCTTGGCAAGGAATCGCTGTGGGAAGAA 840
Qy 1113 GCTTTGCAATGAT 1126
Db 841 GCTTTGCAATGAT 854

RESULT 4
LOCUS AY105934
DEFINITION Zea mays PCO089998 mRNA sequence.
ACCESSION AY105934
VERSION AY105934.1 GI:21209012
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2279)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
```

TITLE Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.  
 Overgo Probes  
 JOURNAL Mapping Project/DuPont Consensus Sequences for Design of  
 REFERENCE Unpublished (2002)  
 2 (bases 1 to 2279)  
 AUTHORS Coe, E.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA  
 COMMENT If you are interested in getting corresponding physical clones,  
 these are publicly available from ZmDB and may be found by BLAST  
 searching at MSU, [maizemap.org](http://maizemap.org); ZmDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR,  
[www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the  
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
[www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

## FEATURES

source

Location/Qualifiers  
 1..2279  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="MaizeDB:635109"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/DuPont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 contigs to seed DuPont contigs; this resource was  
 assembled by DuPont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"

## ORIGIN

Query Match 26.7%; Score 653.8; DB 3; Length 2279;  
 Best Local Similarity 60.3%; Pred. No. 1,7e-159;  
 Matches 1119; Conservative 0; Mismatches 727; Indels 9; Gaps 2;  
 121 AGTGCACCAAGTCAGGTTCCACGCCACACGCGTCTTTTCAAGTCTCTAAGTACTCTTT 180  
 Db AGTGCAACAGGTGGCGCGCCACCGCGCGGAGACAGCGGAGCAAGATGAAGTGAGGGT 247  
 181 GAAGGAGACTTCTTCCCTGATGACCCCTTTGAGCGAGTTCAAGAACAAAGCAGCTTCCAA 240  
 Db GAAGGAGACCTTCTTCCCGAGACCGCTTCGGGCGTTCAAGGGCAGCGCGGGGAC 307  
 241 GAAGTTCAAGCTTGGCCTTCAGTTCCTTCCCATTTTCGAATGGCTCCCAAAATACAC 300  
 Db GCAGTGGCTCATGCGGTGAGTACCTCTTCCCATCTCTGGACTGGGTGCGAGCTACTC 367  
 301 CTTTCAGTTCTTGAAGCTGACCTCATAGTGGGATCACCATCGCTAGCTTGGCCATTCC 360  
 Db CTTTGTGCTCTTCAAGTCCGACCTCGTGGGCGCTCACATTTGCCAGCCTCGCCATTCC 427  
 361 TCAGGGGATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTTGGACTATATTCGAG 420  
 Db TCAGGGCATTTAGCTACCGGAGCTGGCAAGCTTGGCTTCCCTCCATTAATCGGCTGTATTCGAG 487  
 421 CTTTATACCAACCATGATTATCGATGATGGTAGCTCGAGGGATTTGGCAGTGGGAC 480  
 Db CTTTGGTCCCGCATGGTGTACGCGGTGCTGGGAGCTCCGCTGACCTGGCGGTGGGCC 547  
 481 TGTGGCGGTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCGTGATCCCAATGA 540  
 Db GGTGTGATTCGTGCGTGTATCATGGGGTCCATGCTGCGGAGCGGCGTGGAGCCCAACCGC 607  
 541 AGACCCAAAGCTTTACTCTCCACTGGCTTTACAGCTTACATATTTTGTGCTGGTGTTC 600  
 Db GGAGCCGACGCTGTTCTTGAGCTGGCTTCACTCCACCTGTTTGGCGGGCTGGTGCA 667  
 601 GGCTGCTCGGCTTGTGTTTGGTGGGTTGATCGTGGATTTTCTGTCAATCCAAACCAT 660  
 Db GGCTTCCCTGGGCATCTCAGGCTCGGCTTCGTTCATCGATCTTCTGTGCCAAGCGGCT 727  
 661 AATAGGGTTCATGGGAGGACGCCACGGTGGTGTGTCTCGCAACATAAAATCGATTCT 720

728 GGTGGGGTTCATGCGCGGCGCCCATCATCTGTTGGCGCTGCAGCAGCTCAAGCGCTGCT 787  
 721 TGGCCTTTGAGCAATTTCAACCATGGAGCTGATATCATATCAGTGATGGCTCTCTTTTTCAC 780  
 788 GGGCATGCTCCACTTCAACCAACAGATGGCATCGTCCAGTCATGGCTCCGCTCTTCCA 847  
 781 CMAAATCATGAGTGGAGTGGGAAAGTGTGTGTTAGGATGTCTTCAATTTTCTTCT 840  
 848 CCACACAGCGAGTGGTCTGGGAGACGATCTCTCATGGGCGTCTGCTCTCTCTTCT 907  
 841 CTTTACCAACAAGTACTTTCAGCAAAAACGACCAAGGTTTTTTTGGTGTGTCAGCAATGGC 900  
 908 GCTGTGGGAGGATGTGAGCATCAGATGGCCAAAGCTTTTCTGGTTCGGCGTGGCG 967  
 901 GCCATTTGACGTCGTTTATTTGGGAAAGTCTCTTTGGTTTATTTTCACTCACGCCGAGAGCA 960  
 968 GCCCTGGCATCGGTCAACCATCTCGAGCTGCTGTTTCTCTTCAAGCTCAGAACCA 1027  
 961 CGGTGTTGAGTATAGGAGAACTGAAGAGGGTTTGAATCCACCATCACTCAAAATCT 1020  
 1028 TGGCATCAGCATCATTTGGGAGCTCAAGTGGCGCTGAATCGCCCTCGTGGGACAAGCT 1087  
 1021 GGTATTTGTGCGCTTACATGACTACAGCTGTCAAACTGGCATGTGTTGGCATCAT 1080  
 1088 CTTGTTTGACCGCGCTATTTAGGCTCACCATGAAGACTGGGCTTTGTCAACGGAATCAT 1147  
 1081 ATCAGCTTGGGAGGAAATAGCAGTAGGAGAAAGCTTTTGCAAATGTATAAAATTTACAATAT 1140  
 1148 CTCACCTGACGGAAGGAATAGCGGTTGGTAGACATTTGCTCACTCAAGGACTACCGAT 1207  
 1141 TGATGCCAACAAAGAGATGATAGCTATTTGGGACCATGAACGTAGTTGGTCTTTTCACTC 1200  
 1208 AGATGGAAACAAGGAGATGATGGCCATAGGTTGATGAATGTTTGGGTCTTGCAATC 1267  
 1201 TTGCTACCTCACAAAGGACCACTTTTCGGCTGCGCTGGAACATAAAGCTGGATGCAA 1260  
 1268 ATGCTACGTAAACAAGGTGGTCTCCGCTCTGCTGTAACCAACAGCCGCGCTGCAA 1327  
 1261 GACAGCAGCTTCCAAACATTAATGTCACTTTGAGTAATTTTGACATTTGTTATTTCTGAC 1320  
 1328 GACTGCCATGTCCAAGCTGATCATGCGCTGACTGTGATGGTCACGCTGCTGTTCTCAT 1387  
 1321 ACCCTTTTCCATATACATCCCTGGTGGTGTATCAGCTATTTATCGTATCTGCAATGCT 1380  
 1388 GCCATGTTTCTGTGTACACACCAACGTTGCTCTCGGAGCGATCATCATCGCCGCTGAT 1447  
 1381 TGGACTCATAGATTATGAAGAGCCCATCCATCTATTTTAAAGTTGACAAATTTGACTTTGT 1440  
 1448 CGGCTGTATCGATTTCCCGCGGTGTACACATCTGGAAGATGGAACAAGATGATTTCT 1507  
 1441 GGTGTGCATGAGTGCAATTAATTGGCGTGTCTTTTGGCAGTGTGAAATTTGGCTTAGTCAT 1500  
 1508 GGTGTGCTTTGGCGTGTGGCGCTCATCTTCACTCAGTCCAAGAGGCTTGGAT 1567  
 1501 AGCTATGTAATATCTGTACTTGGGTACTTCTATTTTATTTGCAAGGCCAAGGACTTCGT 1560  
 1568 AGCGTGTGTTATCTATATTTAGGGTGTGATGCAAGTCAACAGGCCAAGATGATGTT 1627  
 1561 TTTGGGCAACATTTCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAA 1620  
 1628 TCAAGGGAACATCAAGGGGACTGATATTTACAGAGACCTGTCATCTATCAAGAGGCGCCA 1687  
 1621 ACATGTTCTTGGAAATGCTAAATTTAGAGATTTGATGCAACCAATTTACTTTGCCAATGCCAG 1680  
 1688 AAGAGTTTCTGGTCTTGTGATCTTGGCCATTTGAAGCACCGATAAATCTTGGCAACTCCAA 1747  
 1681 CTATTTAAGAGAAAGATCAAGGTGGATTTGATGAAGAGAGAAAGAAATTTAAAGCTAC 1740  
 1748 CTACCTGTAATGAAGGATTTAAAGATGGAT-----AGAGGAAGAAATCTTTTGAACAGGA 1801  
 1741 AGGGGAGACTAGTTTGCAGTATGTTTATATGATATGATGCTGTTGGAAACATTCATAC 1800

Db 1802 TAAACACTACTGAACCTCCATTTTCATAAATCTTGGATCTGTCTGAGCTGTCTTCCTGCAATTCGACAC 1861

QY 1801 AAGTGGAAATAGATATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGAGCTACAGCT 1860

Db 1862 AAGTGGCATAGCGTTCTCTATTGACATATAAGAAATCAATAGAAACGTTGCTCGAGCT 1921

QY 1861 TGTTTTGGTCAATCTCTGTAGTGAAGTGAAGAGAACTGAAACAAATCGA---AGTTCCA 1917

Db 1922 TGTGCTTGTCAATCCAACTGGGAAGTCAITGGAGAAATACACGTTGCAACGAGGCTGA 1981

QY 1918 AATCATTTTAGGAAGAATAGATCTATCTGATCTGTGAAGAGCGCGTTGGAGCA 1972

Db 1982 AAATATTTTAGGCCAGATTGCTTGTATCTGACCACTGGCGAAGCAATGCGTTCA 2036

RESULT 5

CL959635

LOCUS

DEFINITION OsiFCC003265 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.

ACCESSION CL959635

VERSION

KEYWORDS

SOURCE CL959635.1 GI:52373944

ORGANISM

Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 2001)

AUTHORS Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.

TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis

JOURNAL Unpublished (2004)

COMMENT Contact: Chen Chen  
Department of Bioinformatic  
Beijing Institute of Genomics  
Chinese Academy of Sciences, Beijing 101300, China  
Tel: 86-10-80481559  
Fax: 86-10-80488676  
Email: chenchen@genomics.org.cn  
Rice genomic sequence.  
Class: exon-trapped.

FEATURES

source

1..2001

Location/Qualifiers

/organism="Oryza sativa (indica cultivar-group)"

/mol\_type="genomic DNA"

/db\_xref="taxon:39946"

/clone\_lib="Oryza sativa Express Library"

/notes="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 24.7%; Score 605.4; DB 9; Length 2001;

Best Local Similarity 58.4%; Pred. No. 7.6e-147;

Matches 1097; Conservative 0; Mismatches 776; Indels 6; Gaps 2;

QY 145 GCACAGCGCTTTTCAAGTCTCTAAGTACTCTTTGAAGGAGACTTCTCCCTGATCA 204

Db 123 GCGCCCGCGCTTCGGCGAAGCTGTGTGCGACCTCGCGGAGAGCTTCTTCCCGACGA 182

QY 205 CCCTTTTGGCGAGTTTCAAGAAACAGCCAGCTTCCAGAAAGTTTCATCTTGGCTTCAGTT 264

Db 183 CCGGTTCCGCGGTTCCGCGCGTCCCGCGCCAGCGCGGTGTGCGCGTCAAGTA 242

QY 265 CTTCTTCCCAATTTTCAATGGGCTCCCAAAATACACTTTTCAAGTTCTTGAAGAGTGAACCT 324

Db 243 CTTTCGTCGCGCGCTGGAGCTGGGTCCCGGCTACGGCTCGACAAAGTTCAAGTTTCAGCT 302

QY 325 CATAGCTGSCATCACCATCGTAGCTTGGCATTCTCAGGGCATCAGTTATGCCAAGCT 384

Db 303 CCTCGCGGCGATCACCATCGCCAGCCTCGCCATCCCGCAGGGGATCAGCTACGCCGCGCT 362

QY 385 CGCCAACTCCCTCCAATTTCTTGGACTATATTCGAGCTTTATACCAACCAATTTGATTTATGC 444

Db 363 CGCCAACTCCCTCCAATTCATCGGCTCTATTCGAGCTTCGTGCGCCGCTGATGTACGC 422

QY 445 GATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACTGTGGCGGTTGGATTCGTTCTGAT 504

Db 423 CGTGTTCGGGAGCTCCAAACAACCTCGCGTGGGACGCTGGCGGGCGCTCGCTGTGCT 482

QY 505 GGGTTTCGATGTTGAGTGAATCCGCTGATCCCAATGAAGACCCAAAGCTTTTACCTCCACT 564

Db 483 GCGGCTCCATCATCGAGACGGAGGTGGCGGCGACAGAAACCCGAGCTGTGACTCGCAGCT 542

QY 565 GCGCTTTCACAGCTACATTTATTTGCTGGTGTTTTTTTCAGGCTGCTTGGGCTCTGTTAGGTT 624

Db 543 CTTCTACACCGCGCTTCTTCAACCGGCTCTTCCAGACCGCGCTAGGCGTCTTTAGGTT 602

QY 625 GGGGTTGATCGTGGATTTTCTGTCAATGCAACCAATATAGGGTTTCATGGGAGGAGCAGC 684

Db 603 AGGGCTAATAGTGGATTTTCTCTCGCGGTGCGAGATACCGGTTTCATGGGCGGACGCGC 662

QY 685 CACGGTGGTGTCTCTGCAGCAACTAAATTCGATTTCTTGGCCTTGAGCATTTTCAACCATGG 744

Db 663 GATGATCATCATCATATGTCAGCAGTTCAAGGGCTCTCTGGGATGAAGACATTCACCAACAA 722

QY 745 AGCTGATATCATATCAGTGAATCGCTCTGTCTTTCACCCAAACTCATGAGTGGAGGTGGA 804

Db 723 GACAGACATCATCTCGTCTCTCACTCCACCTTACCATTTACCGGACGAGTGGAAATGGCA 782

QY 805 AAGTGTGTGTAGGATGTCTTCAATTTTCTTCTCTCTAGCACAAAGATCTTACAGCAA 864

Db 783 GAGTGCAGTTCTCGGCATATGCTTCTCTTGTCTGTCTGCTGAGTAAGCAGTCTGAGAAA 842

QY 865 AAAACGACCAAGGTTTCTTGGGTGTCAGCAATGCGCATTTGAGCTCCGTTATATTGG 924

Db 843 GAAACTGCCAAAGTTGTTCTGGGTGTGCGCCATTCGCCCATTCATGTTGTGCTATTCGG 902

QY 925 AAGTCTCTTGGTTTATTTTCACTCACGCGGAGAGCACGCTGTGTAAGTGATAGGAGAACT 984

Db 903 AGGCATCTTTCGCTTCTCTCTCAAGGGTGAACGAGCATGGAATTCGATAGTCGGTAACT 962

QY 985 GAAGAAGGTTTGAATCCACCATCACTACAAATCTGGTATTTGTGTCGCTTACATGAC 1044

Db 963 GAATAAAGGGATCAATCCCTCTGTCATTTGCGCAGTAACTTCCAGTCAGGACATGAA 1022

QY 1045 TACAGCTGTCAAAACTGCGCATTTGCTGGCATCATATCACTTGGGAGGAATAGCAGT 1104

Db 1023 GACGCGAGTGAAGCTGGTCTCTTCTGCGGATCTTAGCACTGGCAGAGGAATAGCGT 1082

QY 1105 AGGAAGAAGCTTTGCAATGTATPAAAAATTTACAATATTGATGGCAAACAAAGAGATGATAGC 1164

Db 1083 CCGCCGAACTTGGCGATGGTCAAGAACGAGCAGATCGACGGCAAACAGGAGATGATCGC 1142

QY 1165 TATTGGGACCATGAACGTAGTTGGTTCTTTTCACTCTTGTCTACCTTCACAAAGGACCAT 1224

Db 1143 GTTTGGTATCATGAACAATCGCGGTTCTTTTCACTCTCTGCTACCTCAGCACAGGCGCGTT 1202

QY 1225 TFCGCTTCGCGTGTGAATATAACGCTGGATGCAAGACAGCAGCTTCCAAACATTATAT 1284

Db 1203 CTCAAAGTCGCGCGTGAATCTCCACCGCGGTGCAAGACCGCGATGTGCAAGCTGGTGT 1262

QY 1285 GTCACCTTCGAGTAATGTTGACATTTGTTATTCCTGACACCCCTTGTTCATTTACACTCCCT 1344

Db 1263 GTCGCTGTGCATCATGCTGGTGTACTGTCTTCTGGCAACCCCTCTTCAAGTACACCCGCT 1322

QY 1345 GGTGTGTATCAGCTATATTCGTATCTGCAATGCTTGGACTCATAGATATTGAAGCAGC 1404

Db 1323 GGTGCGCTCTCTCCATCATCTGTCGCCATGATCGGGCTCGTCAAGGTCAGAGGTT 1382

QY 1405 CATTCATCTATTTAAGGTTGACAAATTTGACTTTTGTGGTGTGCTGAGTGCATATATTGG 1464

Db 1383 CTGCCACCTCTTACAGGTGCAACAGTTGACTTTCGCACTGTCATGGTGGCTTCTCTCGG 1442

QY 1465 CGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTTCATAGCTATTGTAATATCTGTACTTCG 1524



1443 CGTCTCTTCTTACCAATGTCGAGGCTCAGCGCATCGTAGGCTTGCAATGTCAG 1502  
1525 GGTAATCTTATTTATTCGAAGCCCAAGGACATTCGTTTGGGCAACATTCACAAATTCGT 1584  
1503 GGCCTGCTGTACGTGCGGAGGCTGCACCTGCAAGCTCGGAAACATAGCAGGAGCGA 1562  
1585 GATATACGGAATGTTGAGGACATCAAAATGCAAAACATGTTCTCGAATGCTAATCT 1644  
1563 GACCTCCGCGACGTGAAGCAATACCCCAAGCGAAGAGCATCCCTGGCATCTCTCGTCT 1622  
1645 AGAGATTGATGACCAATTTACTTTGGCAATGCCAGCTATTTAAGAGAAAGGATCAACAG 1704  
1623 CGAGCTGGCTCTCCATCTACTTCGTAATGCGGGTTACTGCGAAGAAAGATTTTGAG 1682  
1705 GTGGATTGATGAAGAAGAAAGAAATTAAGCTTACAGGGGAGACTAGTTTGCAGTATGT 1764  
1683 ATGG--GTGGAAGATGAGGATAACCTGTGCAAGAGCGTCGGGCACGATTTGCAATACTT 1739  
1765 TATTAATGATAGTACTGCTGTGGAACATTTGATACAAAGTGGAAATAGTATGCTTGAGA 1824  
1740 GGTCTTGTGATCTTGGTGTGTCACCTCTGTGCAAACTCTGGAGTTGGGATGCTACTAGA 1799  
1825 GTGAAGAAGATTACAGAGAGAGAGACTACAGCTTGTGTTTGTCAATCCTGTAAGTGA 1884  
1800 AGTACACAGAGCTCGAAGAGAGGATCACGATAGCTCTGACGAATCCGAGGCTAGA 1859  
1885 AGTGATGAAGAACTGAACAAATC--GAAAGTTCCAAATCAATTTAGGGAAGAAATGGAT 1941  
1860 GGTGACAGAGAGCTGTGCTGTGATACGTCAGGAGCATCTTAGGGATGAGTGGGT 1919  
1942 CTATCTGACTGTGGAAGAGCGGTTGGAGATGCAACTTCAATCTAGCTGCAAGCAAAAC 2001  
1920 CTTCTGACGGTCAAGAGCGCATCACGGGTGTCGATACGGCTGCAGATATCCAGAAA 1979  
2002 GAACCCAAAGAAAGATCAA 2020  
1980 TAGGGAGAAGACNAAGTA 1998

RESULT 6  
AW598447  
LOCUS  
DEFINITION BJ91904.y1 Gm-cl023 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl023-1999 5' similar to TR:064435 064435 SULFATE TRANSPORTER.  
[2] TR:065025 ; mRNA sequence.

AW598447  
AW598447.1 GI:7285969

EST.  
Glycine max (soybean)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 583)

Shoemaker, R., Keim, P., Vodkin, L., Erpel, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, J., Jackson, J., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)

Other ESTs: BUS44910 corresponding to Gm-r1088-1045 (3')  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this

clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com) Insert Length: 1570 Std Error: 0.00 High quality sequence stop: 440.

#### FEATURES

source

Location/Qualifiers

1..583

/organism="Glycine max"

/mol\_type="mRNA"

/cultivar="T157"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl023-1999"

/tissue\_type="seed coats of greenhouse grown plants"

/lab\_host="DH108"

/clone\_lib="Gm-cl023"

/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This cDNA library was constructed from mRNA isolated from seed coats (100-200mg) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restriction site. Sal I linkers adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

#### ORIGIN

Query Match 23.7%; Score 581.4; DB 2; Length 583;  
Best Local Similarity 99.8%; Pred. No. 1e-140;  
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1198 CTCTTGCTACCTCACACAGGACCAATTTGCGCTTCGGCTGCTCACTATCAACCTGATG 1257  
Db 1 CTCTTGCTACCTCACACAGGACCAATTTGCGCTTCGGCTGCTCACTATCAACCTGATG 60  
Qy 1258 CAAGACAGCAGCTTCCAAACATTAATATGTCACCTTGCAGTAATTTGACATTTGTTTCCT 1317  
Db 61 CAAGACAGCAGCTTCCAAACATTAATATGTCACCTTGCAGTAATTTGACATTTGTTTCCT 120  
Qy 1318 GACACCTTGTTCATTAACACTCCCTGGTGGTCTATCAGCTATTTATCGTATCTGCAAT 1377  
Db 121 GACACCTTGTTCATTAACACTCCCTGGTGGTCTATCAGCTATTTATCGTATCTGCAAT 180  
Qy 1378 GCTTGACATCATAGATTATGAGAGCCCATCCATCTATTAAAGTTTGACAAATTTGACTT 1437  
Db 181 GCTTGACATCATAGATTATGAGAGCCCATCCATCTATTAAAGTTTGACAAATTTGACTT 240  
Qy 1438 TGTGGTGTGCATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTTCGAAATTTGGCTTAGT 1497  
Db 241 TGTGGTGTGCATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTTCGAAATTTGGCTTAGT 300  
Qy 1498 CATAGCTATTGTAATATCTGTACTTCGGGTACTCTTCTATTATTGCAAGGCCAAGGACATT 1557  
Db 301 CATAGCTATTGTAATATCTGTACTTCGGGTACTCTTCTATTATTGCAAGGCCAAGGACATT 360  
Qy 1558 CGTTTGGGCAACATTTCCAAATCTCTGATATACCGAAATGTTGAGCAGCTATCAAAATGC 1617  
Db 361 CGTTTGGGCAACATTTCCAAATCTCTGATATACCGAAATGTTGAGCAGCTATCAAAATGC 420  
Qy 1618 AAAACATGTTCTCGGAATGCTAATCTAGAGATTGATGACCAATTTTACTTTGCCAATGC 1677  
Db 421 AAAACATGTTCTCGGAATGCTAATCTAGAGATTGATGACCAATTTTACTTTGCCAATGC 480  
Qy 1678 CAGCTATTTAAGAGAAAGGATCAACAGGTGGATTTGATGAAGAGAGAAAGAAATTAAGC 1737  
Db 481 CAGCTATTTAAGAGAGAAAGATCAACAGGTGGATTTGATGAAGAGAGAAAGAAATTAAGC 540  
Qy 1738 TACAGGGGAGACTAGTTTTCAGTATCTTATTTATGATATGAGT 1780



Db	541	TACAGGGGAGACTAGTTTTCAGTATGTTATTAATGATAGT	583	QY	784	AACTCATGATGGAGTGGAAAGTCTGTGTTAGGATGTGCTTTCATTTTCTTCTCT	843
RESULT 7				Db	594	CACCAAGGAGTGGTGGTGGACGATCCTGATGGCCGTGTCTTCTGCTCTCTGCT	653
CL969864				QY	844	TAGCA CAAGATACTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTACGAATGGCGCC	903
LOCUS		1812 bp DNA linear GSS 21-SEP-2004		Db	654	CACGGCGAGGCATGTGAGTATGAATGGCAAAAGCTTTTCTGGGTTTACGTTGTGCTCC	713
DEFINITION		Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.		QY	904	ATTGACGTCCTGTTATATTGGGAAGTCTCTTGGTTTATTTTCACTCAGCGCGAAGACAGG	963
ACCESSION	CL969864	GI:52394347		Db	714	CTTGGCATGTGTTATCTCTACCCCTGCTTGTTCCTGTTCAAGCTCAGAAACATGG	773
VERSION	CL969864.1			QY	964	TGTTGAAGTGATAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAAATCTGGT	1023
KEYWORDS	GSS.			Db	774	CATTAGCATAATTGGGCAGCTCAATGTGGCCTGAATGCCCTTCATGGGATAAATACT	833
SOURCE	Oryza sativa (indica cultivar-group)			QY	1024	ATTGTGTCGCTTACATGACTACAGCTGTCAAAACTGTCATGTGCTGGCATCATATTC	1083
ORGANISM	Oryza sativa (indica cultivar-group)			Db	834	CTTTGATCCTCAATATTGGGCTCTCACTGTAAAGACTGGCCTTGTCAACGGAATCATCTC	893
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.			QY	1084	ACTTGGGAAGCAATAGCAGTAGGAGAGCTTTTCCAAATGTATAAAATTTCAATATTCGA	1143
REFERENCE	1 (bases 1 to 1812)			Db	894	CCTAACGGAAGGAGTGGGAGTGGTTCCTCACTCAAGGACTACAGGTAGA	953
AUTHORS	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.			QY	1144	TGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTCACCTCTTG	1203
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis			Db	954	TGGTAACAAAGAGATGATGGCCATAGGATTGATGAACATTTGGGCTCATGTACATCATG	1013
JOURNAL	Unpublished (2004)			QY	1204	CTACCTCAACAAGGACATTTTCGGTTCGGCTGTGAACACTATACGCTGGATGCAAGAC	1263
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped.			Db	1014	CTATGTAACTACAGGAGCATTTCTCTGTTCTGTGAACCAATGCTGTTGCAAGAC	1073
FEATURES	Location/Qualifiers			QY	1264	ACGAGCTTCCAACTTATTAATGTCATCTTGCATTAATGTTGACATTTGATTTCTTCACACC	1323
source	1. .1812			Db	1074	TGCCATGTCAAATGTGATCATGGCACTGACTGTCAATGTTGACACTGTTTCTCATGCTC	1133
	/organism="Oryza sativa (indica cultivar-group)"			QY	1324	CTTGTTCATTAACACTCCCTCGTGTGCTATTCAGCTATTTATCGTATCTGCAATGCTGG	1383
	/mol_type="genomic DNA"			Db	1134	ACTGTTTGTGTACACACCGAATGTTGTTCTTGGAGCAATCATTAATGCTGCTGTGATTGG	1193
	/db_xref="taxon:39946"			QY	1384	ACTCATGATTAATGAAGACCATCCATCTATTTAAGGTTGAACAAATTTGACTTTGCTGTGTT	1443
	/clone_lib="Oryza sativa Express Library"			Db	1194	GCTCATGATCTCCCTGCTGTATACAACATCTGGAAGATGGAACAAGATGGAATTTCTGTT	1253
	/note="Oryza sativa exon trapped genomic sequences"			QY	1444	GTGCATGATGTCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAATAGC	1503
Query Match	23.3%; Score 570.8; DB 9; Length 1812;			Db	1254	GTGCTCATGTGATTTGCTGTTGTTAATCTTCACTCAGTCCAAAGGCTTGCAATAGC	1313
Best Local Similarity	61.0%; Pred. No. 8.5e-138;			QY	1504	TATTTGTAATATCTGTAATTTCTGGGTACTTCTATTTTATTTGAAGCCCAAGGACATTCGTTTT	1563
Matches	945; Conservative 0; Mismatches 597; Indels 6; Gaps 1;			Db	1314	GTTTGGTATATCAATATTTCAGGGTACTGTTGCAAAATCAAGGCCAAAGATGATTCTCA	1373
364	GGGCATCAGTTAGCCAAAGCTGCCAACTCCCTCCAAATCTTGGACTATATTCGAGCTT	423		QY	1564	GGGCAACATTCOAATTTCTGTGATATACCGAAATGTTGAGGACTATCAAAATGCAAAACA	1623
174	GGGCATTAGCTACGGAAGCTGGCGAGCTTGCCTCCGATCATCGGCTCTATTCGAGCTT	233		Db	1374	AGGGAACATCAAGGCACTGATATTTACCGAAATCTTTCATCAGTACAGGACGCTCAAG	1433
424	TATACCACCATTTGATTTATGCGATGATGGTAGCTCGAGGATTTGGCAGTGGGACTGT	483		QY	1624	TGTTCTCGAATGCTAAATTTCTAGAGATGATGACCAAAATTTTACTTTTGCCTGCGAGCTA	1683
234	TGTGCCCGCATGTTGTACGCGTGTGGGAGCTCGAGGACCTGGCGGTGGGCGCGT	293		Db	1434	AGTCTCGGTTCTTAATCTGACAGTGGAGCTCTTATAAACTTTGCAACACCAACTTA	1493
484	GGCGTTTGGATCGCTTCTGATGAGGTTGATGTTGAGTAATGCGTTGATCCCAATGAAGA	543		QY	1684	TTTAAAGAGAAAGATCACAAAGTGGATGATGAAGAAGAAAGAAATTTAAAGCTACAGG	1743
294	GTGATCGCGTCTGCTGATATGCGGTCCATGCTGCGCAGCGCGGTGAGCCCGCCGCA	353		Db	1494	CCTAAATGAAGGATTTAAAGATGAT-----AGAAGAGAAAGTCTTCGAGGAACCAA	1547
544	CCCAAGCTTTACCTCCACTGCTTTACAGCTACATTTATGCTGTTGTTTTCAGGC	603		QY	1744	GGAGACTAGTTTGGAGTATGTTTATAATGATGATGCTGTTGGAAAACATTTGATACAG	1803
354	GCGCTGCTCTTCTCCAGCTCGCTTCCACTTCCCTTCTTGGCGGCTCGTCCAGGC	413		Db	1548	ACAAAGTGAATCCATTTCTGTAATTTTGGATCTGTGAGCTGCTCCCTGCAATTTGATACAAG	1607
604	TGCTTTGGGTCTGTTTATGTTGGGTTGATCTGGGATTTTCTGTACATGCAACCAATA	663		QY	1804	TGGAATAAGTATGCTTGAAGAGGTCGAAGAAGATTACAGAGAGAAGAGCTACAGCTCT	1863
414	CTCCCTCGGCATCTCCGGCTGGGTTTCATCATCATGATTTCTGTGCAAGCGCAGCTTGT	473		Db	1608	TGGCATATGTTCTTCATCGACTTGAAGAAATCAACAGAGAAGCATGAGCTAGAGCTTAT	1667
664	AGGTTTCATGGAGAGCAGCCACGCTGTTGTTCTGACGAACTAAATTCGATTTCTGG	723		QY	1864	TTTGTGTCATCTGTAAAGTGAAGTGAAGAAACTGAACAAATCGAA	1911
474	CGGTTTCATGGCGCGCGCGCATCATCTGTTGCTGCGTGCAGCAGCTCAAGGCGCTGCTCG	533		Db			
724	CCTTGAGCAATTTCAACCATGGAGCTGATATCATATCAGTATGATGCTCTGTTTTCACCA	783		QY			
534	CATCGTCCACTTACCACCCAGATGGGCTCGTCCCGCTCATGGCTCCGTCATCCACCA	593		Db			



FEATURES	High quality sequence stop: 394.		KEYWORDS	EST.		
	source	Location/Qualifiers				
ORIGIN	1. 522		SOURCE	Glycine max (soybean)		
	/organism="Glycine max"					
	/mol_type="mRNA"					
	/cultivar="Corolla"					
	/db_xref="taxon:3847"					
	/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-4125"					
	/tissue_type="floral meristematic mRNA"					
	/lab_host="DH108"					
	/clone_lib="Gm-cl051"					
	/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from floral meristematic mRNA provided by Dr. Halina Knap of Clemson University. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH108 host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."					
	Query Match 21.2%; Score 520.4; DB 4; Length 522; Best Local Similarity 99.8%; Pred. No. 9.4e-125; Matches 521; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
	Qy	1868			GTCAATCCTGTAAGTCAGATGAGAACTGAACAAATCGAAGTTCGAAATCATTTA	1927
Db	1	GGCAATCCTGTAAGTCAGATGAGAACTGAACAAATCGAAGTTCGAAATCATTTA	60			
Qy	1928	GGCAAGAAATGGATCTATCTGACTGTTGAAGAGCGCTTGGACATCGCACTTCAATCTA	1987			
Db	61	GGGAGAAATGGATCTATCTGACTGTTGAAGAGCGCTTGGACATCGCACTTCAATCTA	120			
Qy	1988	CGTGAAGCAAAAACGAAACCAAGAAAGATGAAACAGAGGGTTGGAACAATGTGCACTG	2047			
Db	121	CGTGAAGCAAAAACGAAACCAAGAAAGATGAAACAGAGGGTTGGAACAATGTGCACTG	180			
Qy	2048	AGTCATATGCCAAGATATCTTAATAACTCAAAAAGCTTATTCGTTTCGCTTAGTA	2107			
Db	181	AGTCATATGCCAAGATATCTTAATAATACTCAAAAAGCTTATTCGTTTCGCTTAGTA	240			
Qy	2108	ATGTTACCACTACAATGTGGCATGAGAAATTTCTGAATCACGCCGAAGATTTTAAAG	2167			
Db	241	ATGTTACCACTACAATGTGGCATGAGAAATTTCTGAATCACGCCGAAGATTTTAAAG	300			
Qy	2168	GCATAGGAAAATGAAAGATGCAAGGGTCTTCTAATTTCTCAACTCTGCATCCTTAGTAG	2227			
Db	301	GCATAGGAAAATGAAAGATGCAAGGGTCTTCTAATTTCTCAACTCTGCATCCTTAGTAG	360			
Qy	2228	AAGAAATCTCCTATGATAGGCTGTGAAATTAATCTTTACGTATCATCTCGTAATAT	2287			
Db	361	AAGAAATCTCCTATGATAGGCTGTGAAATTAATCTTTACGTATCATCTCGTAATAT	420			
Qy	2288	ATTCAAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACACATGTAAGAG	2347			
Db	421	ATTCAAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACACATGTAAGAG	480			
Qy	2348	GTAAGAAGTGTGCTAGCACTCTCCCATATTCATTTGTA	2389			
Db	481	GTAAGAAGTGTGCTAGCACTCTCCCATATTCATTTGTA	522			
RESULT 10						
CO982632/c						
LOCUS						
DEFINITION						
CO982632						
ACCESSION						
VERSION						
CO982632.1 GI:51341315						
REFERENCE						
AUTHORS						
Vodkin, L., Shoemaker, R., Keim, P., Polacco, J. C., Retzel, E., Khanna, A., Shealy, R., Clough, S., Thibaud-Nissen, F., Coryell, V., Erpelnding, J., Gonzalez, D. O., Stromvik, M., Rodriguez-Huete, A. M., Schweizer, P., Gong, G. and Liu, L.						
A Functional Genomics Program for Soybean (NSF 9872565) (2004)						
Unpublished (2004)						
Other ESTs: BU090662 corresponding to Gm-cl066-830 (5')						
Contact: Vodkin, L. O., PI, A Functional Genomics Program for Soybean (NSF 9872565)						
Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics						
University of Illinois						
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA						
Tel: (217) 244-6147						
Fax: (217) 333-4582						
Email: l-vodkin@uiuc.edu						
Plate: GM89016A2 row: D column: 09						
Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'						
High quality sequence stop: 594.						
FEATURES						
source						
1. 594						
/organism="Glycine max"						
/mol_type="mRNA"						
/db_xref="taxon:3847"						
/clone="Gm-r1089-5922"						
/clone_lib="Gm-r1089"						
/note="The library Gm-r1089 is a sequence-driven, reracked set of 9,216 low redundancy clones selected from 38 different cDNA libraries constructed from various tissues and stages of development of soybean including 973 cDNAs from germinating cotyledons (source library Gm-cl069, Gm-cl076, and Gm-cl077); 1,465 cDNAs from various tissue and organ systems of the adult plant; 476 cDNAs from adult stem tissue (source library Gm-cl062); 1340 cDNAs from tissue culture derived somatic embryos (source libraries Gm-cl036 and Gm-cl075); 2918 cDNAs from hypocotyls or young seedlings; 742 cDNAs from germinating seedlings, shoot tips, or leaves exposed to various stresses (source libraries Gm-cl065, Gm-cl066, and Gm-cl068); 839 cDNAs from young leaves or hypocotyls exposed to bacterial and fungal pathogens (source libraries Gm-cl072, Gm-cl073, Gm-cl074; and Gm-cl084); and 463 from roots of young plants grown in hydroponic media without phosphate (source library Gm-cl087). The 5' ESTs of the source clones from the different libraries were used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1089 and the cDNA clones of the reracked Gm-r1089 library were then sequenced at the 3' end. The unigene selection and 3' sequencing was funded by NSF Plant Genome project #9872565 (http://soybeanomics.cropsci.uiuc.edu/) as part of creation of a low redundancy soybean cDNA set. The source cDNA libraries were constructed by the laboratories of Lila Vodkin, University of Illinois, Randy Shoemaker, Iowa State University, and Paul Keim, Northern Arizona University as part of the Public EST project. http://129.186.26.94/soybeanest.html. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.ahc.umn.edu/biodata/nsfsoy/. Reracking and 3' sequencing were conducted by services of the University of Illinois Keck Center for Comparative and Functional Genomics http://www.biotech.uiuc.edu/keck.htm. Note: The corresponding 5' EST from each clone in the Gm-r1089						

library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the clone ID of the original cDNA library that is also listed under 'OTHER EST'."

cDNA library that is also listed under OTHER EST :	
ORIGIN	
Query Match Best Local Similarity Matches	21.1%; Score 516.4; DB 7; Length 594; 95.3%; Pred. No. 1.1e-123; 543; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
QY	4 CGAGCTAGCTCCGACATTAAGTTATATACACATAT-TTTGCTTGGCTTAGAAATACTATTGA 62
DB	570 CTAGCTAGCTCCGACATTTAAATATATATACACATATTTTGTAGCCAGAAATACTATCA 511
QY	63 TTGAAGATATGGGAGTGTAGATATATGAGTACCCCTTTGGGCATGAACAACATTTTCAGAGAG 122
DB	510 TTGAATATATGGGAGTGTAGATATATGAGTACCCCTTTTGGGCATGAACAACATTTTCAGAGAG 451
QY	123 TGCACCAAGTCGAGGTTCCACCGCACAGCCGTTTTTCAAGTCTCTTAAAGTACTCTTTTGA 182
DB	450 TGCACCAAGTCGAGGTTCCACCACTCAACCGTTTTTCAAGTCTCTTAAAGTACTCTTTTGA 391
QY	183 AGGAGACTTTCTTCCCTGATGACCCCTTTGAGCGAGTTCAGAAACAAGCCAGCTTCCAGA 242
DB	390 AGGAGACTTTCTTCCAGATGACCCCTTTGAGCGAGTTCAGAAACAAGCCAGCTTCCAGA 331
QY	243 AGTTTCATGCTTGGCCCTTCAGTACTCTTCTTCCCATTTTCCGATTTGGGCTCCCAATACACCT 302
DB	330 AGTTTCATGCTTGGCCCTTCAGTACTCTTCTTCCCATTTCTCGATTTGGGCTCCGATACACCT 271
QY	303 TTCAGTTCTTGAAGCTGACCTCATAGCTGGGCATCACCATCGTAGCTTTGGCCATTTCTCTC 362
DB	270 TTCAGTTCTTGAAGCTGACCTCATAGCTGGGCATCACCATCGTAGCTTTGGCCATTTCTCTC 211
QY	363 AGGGCATCATGTTATGCCAGCTCGCAACCTCCCTCCAAATCTTTGGACTATATTCGAGCT 422
DB	210 AGGGCATCATGTTATGCCAGCTCGCAACCTCCCTCCAAATCTTTGGACTATATTCGAGCT 151
QY	423 TTATACCACCATTTGATTTATGCGATGATGGTGTAGCTCGAGGGATTTTGGAGTGGGACTG 482
DB	150 TTATACCACCATTTGATTTATGCGATGATGGTGTAGCTCGAGGGATTTTGGAGTGGGACTG 91
QY	483 TGGCGTTGGATCGCTTCTGATGGGTTGATGTTGAGTAAATCCGTTGATCCCAATCAAG 542
DB	90 TGGCAGTTGGATCGCTTCTGATGGGTTGATGTTGAGTAAATCCGTTGATCCCAATGAAG 31
QY	543 ACCCAAGCTTTACCTCCACCTGGCTTTCA 572
DB	30 ACCCAAGCTTTACCTCCACCTTGGCTTTCA 1
RESULT 11	
CNS09YVW	2221 bp mRNA linear HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION	GSUFGH70ZF01 of Hormone Treated Callus of strain col-0 of
ACCESSION	Arabidopsis thaliana (thale cress).
VERSION	BX832405.1 GI:42455111
KEYWORDS	HTC; GSU cDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 2221)
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL	Unpublished

Db 701 AGCTTAAGGGCATCTTTGGACTAGTCCATTTTACGCACAAAACCGATGTCGTATCAGTTC 760  
QY 765 TGGCTCTGTTTTACCCAACTCATGAGTGGAGGTGGGAAAGCTGCTGTAGGATG 824  
Db 761 TCACCTCAATCTTAGACAACCGGGCTGAGTGGAAAGTGGCAAGTACCTTGGCTGGTGTCT 820  
QY 825 TCTTCATTTTTCTTCCCTCTTAGCACAGATAC-TTCAGCAAAAACGACCAAGGTTTTTT 883  
Db 821 GCCTTCTGTATTTTACAACTCACTCGATACATAAAGCAAGATACCCCAAGCTATTTT 880  
QY 884 TGGGTGTCAGCAATGGCGCAATTGACGTCGTTATATATGCGAAAGTCTCTTGGTTTATTC 943  
Db 881 GGGGTATCAGCAATGGGTCCAATGGTGTGCTGCTGTTGGTGTGCTGTTGCTTATTTG 940  
QY 944 ACTCAGCCGAGAGACGGTGTGAAGTGATAGGAACTGAAGAAGGTTTTGAATCCA 1003  
Db 941 GTGAAGGAACAGACACGGCACTTCTATCTGTTGGGCTTTAAGAAAGACTAAATCCT 1000  
QY 1004 CCATCACTCAAAATCTGGTATTTGTGTCGCTTACATGACTACAGCTGTCAAACCTGGC 1063  
Db 1001 CCTTCAT-TCACTTTTGAATTTGACTCCAAGTATCTAGGATGGTGTAAAGCGCGGA 1059  
QY 1064 ATTGCTGGGCAATCATACACTTGGGAAGAAATAGCAGTAGGAAGAGCTTTTGAATG 1123  
Db 1060 ATCGTCACAGGACTCATCGCTTTGGCGGAAGGATAGCAATTTGGAAGAAGCTTTGCTGTA 1119  
QY 1124 TATAAAATTTACAAATTTGATGCAACAAAGAGATGATGCTATTTGGACCATGAACGTA 1183  
Db 1120 ATGAAGACGAACAACGGACGGGAACAAGAGATGATGACTTTGGTCTTATGAACGTA 1179  
QY 1184 GTTGGTTCTTTCACTCTTGTCTACCTCAACAAGACCAATTTTCGGGTTCGGTGTGAAC 1243  
Db 1180 ATTGGCTCTTCACTTCTTGTCTATTGACACGGGCCATTTTCAAGACAGCAGTGAAC 1239  
QY 1244 TATAACGCTGGATGCAAGACAGAGCTTCCAAATTAATATGTCATCTTGCAGTAATGTTG 1303  
Db 1240 TACAACGGGGGCAAAAGACACCGATGTGCAATGTAGTAATGGCGCTTTGTCATGATGCTT 1299  
QY 1304 ACATTTGTTATCTCTGACCCCTTTGTCATTACACTCCCTGGTGTGCTATCAGCTATT 1363  
Db 1300 GTGCTTCTTCTTAGGCTCTATTAGCTACACACATAGTGTGCTCTCAGCCATC 1359  
QY 1364 ATCGTATCTGCAATGCTTGGACTCATAGATTATGAAGACGCCATCCATCTATTTAAGTTT 1423  
Db 1360 ATTATGTACGCCATGTTGGGTCTCATCACTACGAGGAGATGACCATCTCTTCAAGTT 1419  
QY 1424 GACAAATTTGACTTTGTTGGTGCATGAGTGATACATTTGGCGTGTCTTTTGGCAGTGT 1483  
Db 1420 GACAAGTTTGATTTCTCTCGTCTGATGTCGCTCTTTTTCGGTGTTCATTCTTTAGCATG 1479  
QY 1484 GAAATTTGGCTTAGTCTAGCTATTGTAATATCTGTACTTTCGGTACTCTTATTTATTGCA 1543  
Db 1480 GACTACGCTCTATCATCTCCGTTGGATTTTCTATCGTAAGACTTTGTTGACGTGTGA 1539  
QY 1544 AGCCCAAGGACATTCGTTTTGGGCAACATTTCCAAATTTCTGTGATACCGAAATTTGAG 1603  
Db 1540 AGSCCTCCACTTGTAATTTGGGAAGGATACCAAACTCGGTTATGTTAGAGACATAGAA 1599  
QY 1604 CACTATCAAAATGCAAAACATGTTCTCTGGAATGCTAATTTCTAGAGATTTGATGCAACAA 1663  
Db 1600 CAGTACCTCGCTCCGAGGAGATGCTTGGTTATATCAATTTCTCAATTAGGCTCTCCAGTC 1659  
QY 1664 TACTTTGGCAATGCCAGCTATTTAAGAAAGGATCACAAGTGGGATTCATGAAGAAGAA 1723  
Db 1660 TTTTTTGGCAATAGCACTTACGTCCTGGAATGCTAATTTCTAGAGATTTGATGCAACAA 1719  
QY 1724 GAAAGAAATTAAGCTACAGGGGAGACTAGTTTCAGTATGTTATATAATTTGATATGAGTGT 1783  
Db 1720 GAAGCCATCG-----AGTTTCTTCTCTGATCTCTCTGGA 1755  
QY 1784 GTTGGAAACATTGATACAGTGAATAGTATGCTTTGAAGAGGTGAAGAAGATTACAGAG 1843

Db 1756 GTTTCAACCACTTGATCATGACCGGATGGAACACACTACTCTGAAATTCAGAGAAATCTCTTGA 1815  
QY 1844 AGAAGAGACTACAGCTTCTTTTGGTCAATCTGTAAGTGAAGTGAAGAACTGAAC 1903  
Db 1816 TCAAAAACATCAAGATGGTGATATAAATCCAGATTTGAAGTCTTGGAAAGATGATG 1875  
QY 1904 AAATCGAATTCCTCAAAATCAATTTAGGGAAGAATGGATCTATCTGACTGTTGAAGAGGCC 1963  
Db 1876 TTGTCCCAATTCGTGGAGAAGATAGGAAAGAGTATATGTTCTTATCGATTCAGATGCA 1935  
QY 1964 GTTGGAGCATCAACTTCAATCTACGTGCAAGCAAAAACGAA 2004  
Db 1936 GTTCAAGCATGCGGATTTAATCTTACCACCAACCAAGCCGGA 1976

RESULT 12  
CV130479  
LOCUS  
DEFINITION  
X9P02a03 Populus stem seasonal library Populus deltoides cDNA, mRNA  
sequence.  
ACCESSION  
CV130479  
VERSION  
CV130479.1 GI:51874399  
KEYWORDS  
EST.  
SOURCE  
Populus deltoides  
ORGANISM  
Populus deltoides  
REFERENCE  
1 (bases 1 to 828)  
Park, S. and Han, K.-H.  
AUTHORS  
Gene expression profile during seasonal growth cycle in poplar tree  
TITLE  
Unpublished (2003)  
JOURNAL  
Department of Forestry  
COMMENT  
Contact: Kyung-Hwan Han  
Michigan State University  
126 Natural Resources, East Lansing, MI 48824-1222, USA  
Tel: 517 353 4751  
Fax: 517 432 1143  
Email: hanky@msu.edu.

FEATURES  
source  
1..828  
Location/Qualifiers  
/organism="Populus deltoides"  
/mol\_type="mRNA"  
/strain="ILL-129"  
/db\_xref="taxon:3696"  
/tissue\_type="stem"  
/dev\_stage="1 year old"  
/clone\_lib="Populus stem seasonal library"

ORIGIN  
Query Match 20.5%; Score 501.4; DB 7; Length 828;  
Best Local Similarity 77.1%; Pred. No. 1e-119;  
Matches 610; Conservative 0; Mismatches 181; Indels 0; Gaps 0;  
QY 1078 CATATCACTTCGGAAGGAATAGCAGTAGGAAGAGCTTTTGCAATGTATAAAATTTACAA 1137  
Db 37 CATTCGGCCGGGAAGGAATAGCAGTAGGAAGAGTTTTCGCATGTTCAAGAAATTTACCA 96  
QY 1138 TATTGATGGCAACAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTTTCAC 1197  
Db 97 TATAGATGGCAACAAGAGATGATGCAATTTGGGACCATGAACATTTAGGGCTTTGCAC 156  
QY 1198 CTCTTGCTACCTCAACAAGAGCAATTTTCGGGTTCGGGTGGAACATAAAGCTGGATG 1257  
Db 157 CTCTTGCTATCTCAACAAGAGCCATTTCTCGGATCGGCTGTGAATTTCAATTCAGGATG 216  
QY 1258 CAAGAAGCAGCTTCCAAATTTAATGTCACTTGCAGTAATGTGACATTTGTTATTCCT 1317  
Db 217 CAAGAAGCAGCTTCCAAATTTAATGTGTCAGCTTGGCAGTCATGGTCACTGTTATTCCT 276  
QY 1318 GACACCTTGTTCATTTACATCCCTCGGTGGTGTATCAGCTATTCATCGTATCTGCAAT 1377  
Db 277 AAGCGCAATGTTCCCAATTTACACTCCACTTTGTGTTCTTATCTCTATTTATTTCTCTCCAT 336





DEFINITION EST715284 potato abiotic stress cdna library Solanum tuberosum cDNA clone POACL46 5' end, mRNA sequence.

ACCESSION CK269206

VERSION CK269206.1 GI:39826184

KEYWORDS EST.

SOURCE Solanum tuberosum (potato)

ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum.

1 (bases 1 to 982)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

Generation of ESTs from abiotic stressed potato tissue

Unpublished (2003)

Other ESTs: EST715285

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>.

Seq primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers

1..982

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"

/clone="POACL46"

/tissue\_types="abiotic stress treated leaf and root tissue"

/lab\_host="DH10B-Tona"

/clone\_lib="potato abiotic stress cDNA library"

/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI, Site 2: NotI;

supplier: Solanum tuberosum var. Kennebec plants were

grown from cuttings on a 16hr light/8 hr dark cycle at 25

C for 3-4 weeks. Abiotic stress conditions were applied to

four separate sets of plants. Set 1 involved saturation of

the soil with 150 mM NaCl and tissues were harvested at

following application of the salt stress (leaves: 2hr,

6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).

Set 2 were grown under the standard conditions and then

were water stressed by withdrawal of further watering

applications. Drought stressed plants were harvested after

cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d

and 5d). Set 3 were grown under the standard conditions

and then were cold stressed by placement at 4 C. Cold

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,

and 4d. Set 4 were grown under the standard conditions and

then were heat stressed by placement at 35 C. Heat

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

2d and 4d and heat-stressed roots were harvested at 6 hr,

12 hr, 1 d, and 4d. RNA was isolated from all tissues and

equal RNA from each tissue and stress was pooled to

construct the cDNA library. RNA sample."

#### ORIGIN

Query Match 19.4%; Score 476; DB 7; Length 982;  
Best Local Similarity 71.5%; Pred. No. 4.8e-113;  
Matches 626; Conservative 0; Mismatches 250; Indels 0; Gaps 0;

QY 123 TGCACCAAGTCGAGGTTCCACCCGCGTTTTCAGTCTCTAAAGTACTCTTTCA 182

DB 101 TACACAGAGTTGAATTCCTCCACACGCGCTTTTTCAGTCAATGAAATACAGTGA 160

QY 183 AGAGACTTTCTCCCTGTATGACCTTTGAGGCGAGTTCAGAACACAGCGCTTCCAAGA 242

DB 161 AGGAACATTAATTTCCAGATGATCCACTTAAGCAATTCAGAACCAAAAACCTTTGAGAA 220

QY 243 AGTTCAATGCTGCGCTTCAGTCTCTCCCAATTTTCGAATGGGCTCCCAATACACCT 302

DB 221 AATTATATCTGGAGTTCAATATTTTTCCAATTTTGAATGGGTTCAAGATACAAAT 280

QY 303 TTCAAGTTCTTTGAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTC 362

DB 281 TTGGTTCTTTCAAGCAGATCTTATGCTGGAATAACATAGTAGTCTTGTCTATTCCIC 340

QY 363 AGGCGATCAGTTATGCGCAAGCTCGCCAACTCCTCCAATTTCTTGAGCATATATTGAGCT 422

DB 341 AAGGAATAAGCTATGCTAAGCTGGGAACTTGCCACCTATACTTGGCTTATATTCAAGTT 400

QY 423 TTATACCAACCATTTGATTTATGCGATGATGCTGAGGATTTGGCAGTGGGACTG 482

DB 401 TTATTTCCACCATTTAGTGACGATTAATGGGAAGTTCAAGAGATTTAGCAGTGGGAACAG 460

QY 483 TGGCGGTTGGATCGCTTCTGATGGGTTGATGTTGAGTAATGCGGTTGATCCCAATGAAG 542

DB 461 TTGCTGTGGATCTCTTCTCATGGCTTCTATGTTGGAGCTGAAGTTAACCTGCTGAAA 520

QY 543 ACCCAAGCTTTTACCTTCCACCTGGCTTTTCCAGCTACATTTTGTGTGTTTTCAGG 602

DB 521 ATCCAACACTTTATCTTCTCATCTTGTCTTCACTGCCACTTCTTCAACCGGTTATTTCGAA 580

QY 603 CTGCGCTTGGGCTGTTTGTAGGTTGGGTTGATCGTGGATTTTCTGTACATGCAACCATAA 662

DB 581 TAGCACTCGGATTTTTCAGGCTGGGATTTATAGTGGATTTTCTATCACATGCAACCAATTG 640

QY 663 TAGGTTTCATGGGAGGAGCAGCGCTGCTGTCTGCAGCAAACTATAATCGATCTTTG 722

DB 641 TGGGATTTATGGAGGAGCAGCTACAGTAGTATGATCTCCAGCAATTAAGGAATATTG 700

QY 723 GCTTGAGCATTTTACCCATGAGGCTGATATCATATCATGATGAGGCTCTGTTTTCACCC 782

DB 701 GTCTTGAAATTTTACTCATGCACTGATGTTGTCTCTCTTACGTTCTGTATTTTTCAC 760

QY 783 AAACATCATGAGTGGAGTGGGAAGTGTCTGTAGGATGTCTTCAATTTTCTTCTCTCC 842

DB 761 AAACATCACCGTGGGATGGGAAGTGTGGGATTTGTTTCTCTTCTTCTTCTTCTTCTGA 820

QY 843 TTAGCACAGATACTTTCAGCAAAAAACGACCAAGGTTTTTTTTGGTGTGAGCAATGGCGC 902

DB 821 TGCTCGCCAAATTTTGTAGCCAGAGAGACCAAAAGCTGTTCTGGGTATCGGCAATGGCGC 880

QY 903 CATTGACGCTCGTTATATTTGGGAAGTCTCTTGGTTTATTTTCACTACGCGGAGAGCAG 962

DB 881 CATTGACGCTCTCATATTTGGGAAGCTATTTCTCGTTTATATAACGATGCTGAAAAACAG 940

QY 963 GTGTTTGAAGTGTAGGAGAACTGAAGAGGGTTTGA 998

DB 941 GTGTTGCTGTATGAGAGCTGAAGAAAGGATTA 976

BU090662 583 bp mRNA linear EST 02-JUL-2004  
su07b07.y1 Gm-cl066 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-cl066-830 5' similar to TR:064435 O64435 SULFATE TRANSPORTER.  
[3] TR:065025 TR:023250 ;, mRNA sequence.

BU090662.1 GI:22540819

EST.

Glycine max (soybean)

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

1 (bases 1 to 583)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,  
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
McCann, R., Waterson, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

TITLE

JOURNAL



COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
When it has been determined, an EST from the other end of this  
clone is listed in the 'Other ESTs on clone' field. This clone is  
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD  
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)  
High quality sequence stop: 420.

FEATURES

Location/Qualifiers

1..583  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultivar="Williams"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl066-830"  
/tissue\_type="Leaf and shoot tip, salt stressed, 2 week  
old seedling"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl066"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from unexpanded leaves and the shoot tips of 2 week old  
seedlings from the cultivar Williams. The 2 week old  
seedlings were salt stressed in a solution of 500mM NaCl  
for 3 days prior to harvesting. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) sequence with a XhoI restriction site. EcoRI  
adapters were ligated to the blunt-ended cDNA fragments  
followed by XhoI digestion. The cDNA fragments were  
directionally cloned into the EcoRI-XhoI restriction site  
of the pBluescript vector. The ligated cDNA fragments were  
transformed into DH10B host cells (GibcoBRL). This library  
was constructed in the laboratory of Dr. Randy  
Shoemaker."

ORIGIN

Query Match 19.1%; Score 468.2; DB 5; Length 583;  
Best Local Similarity 90.3%; Pred. No. 4.5e-111;  
Matches 511; Conservative 0; Mismatches 54; Indels 1; Gaps 1;  
Qy 4 CGAGCTAGCTCGCACATTAAGTTATATAACACATA-TTTCCTTGCTTAGAAATACTATTA 62  
Db 18 CTAGTAGCTCGCACATTTAATTATATTACACATATTTTGTAGCCAAAGAAATACTATCA 77  
Qy 63 TTGAAGATATGGGAGGTAGATTATGAGTACCCCTTTGGGCATGAACAACCTTTGAGAGAG 122  
Db 78 TTGAATATATGGGAGGTAGATTATGAGTACCCCTTTAGGCATGAACAACGTTGAGAGAG 137  
Qy 123 TGCACCAAGTCGAGGTTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTGA 182  
Db 138 TGCACCAAGTGGAGGTTCCACCACTCAACCGTTTTTCAAGTCTCTTAAGTACTCTTTGA 197  
Qy 183 AGGAGACTTTCTTCCCTGATGACCCCTTTGAGGAGGTTCAAGAACAGCCAGCTTCCAAGA 242  
Db 198 AGGAGACTTTCTTCCCATGATGACCCCTTTGAGGAGGTTCAAGAACAGCCAGCTTCCAAGA 257  
Qy 243 AGTTATGCTTGGCTTCAGTCTTCTTCCCAATTTTCGAATGGGCTCCCAATACACCT 302  
Db 258 AGTTATGCTTGGCTTCAGTCTTCTTCCCAATTTTCGAATGGGCTCCCAATACACCT 317  
Qy 303 TTCAAGTTCTTGAAGCTGACCTCATAGCTGCGATCACCCTGCTAGCTTGGCCATTCCTC 362  
Db 318 TTCAAGTTCTTGAAGCTGACCTCATAGCTGCGATCACCCTGCTAGCTTGGCCATTCCTC 377  
Qy 363 AGGCAATCAGTTATGCGCAAGCTCGCCAACTCCCTCCCAATTTCTTGGACTATATTCGAGCT 422  
Db 378 AGGCAATCAGTTATGCGCAAGCTCGCCAACTCCCTCCCAATTTCTTGGACTATATTCGAGCT 437  
Qy 423 TTATACCACCATTTATTTATGCGATGATGGTAGTGGAGGATTTGGCAGTGGGAGCTG 482

Db 438 TTATACCACCATTTGATCTATGCGATGATGGGTAGCTCGAGGGGATTTATGCGAGGGGACTG 497  
Qy 483 TGGCGGTTGGATCGCTTCTGATGGGTTCCGATGTTGAGTAATGCGCTTGTATCCCAATGAAG 542  
Db 498 GGGCAGTTTGAAGGCTTCTTATGGGGTTCGATGTTGAGTATGCGGTCGCTCCCTATGAAG 557  
Qy 543 ACCCAAAGCTTTACCTCCACCTGGCT 568  
Db 558 ACGCCATGCTTTACCTCCACCTTCGCT 583

Search completed: August 31, 2005, 08:45:56  
Job time : 7928 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2005, 03:36:12 ; Search time 122 Seconds

(without alignments)  
2155.714 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGWNV 680

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3475	100.0	680	3 AAY44943	Aay44943 Soybean s
2	2641.5	76.0	658	3 AAG32200	Aag32200 Arabidops
3	2638.5	75.9	658	3 AAG31621	Aag31621 Arabidops
4	2638.5	75.9	703	3 AAG31620	Aag31620 Arabidops
5	2164	62.3	579	3 AAY44936	Aay44936 Corn sulp
6	2070	59.6	510	3 AAG32201	Aag32201 Arabidops
7	2067	59.5	510	3 AAG31622	Aag31622 Arabidops
8	1901	54.7	621	3 AAY44942	Aay44942 Soybean s
9	1898.5	54.6	631	3 AAG329247	Aag329247 Arabidops
10	1855	53.4	457	3 AAG32202	Aag32202 Arabidops
11	1843.5	53.1	688	3 AAY44935	Aay44935 Corn sulp
12	1807.5	52.0	656	3 ABP53777	Abp53777 Arabidops
13	1798.5	51.8	656	3 AAY44944	Aay44944 Wheat sul
14	1444.5	41.6	499	3 AAG329248	Aag329248 Arabidops
15	1271.5	36.6	446	3 AAG329249	Aag329249 Arabidops
16	961.5	27.7	593	3 AAY44945	Aay44945 Wheat sul
17	731.5	21.1	226	3 AAY44938	Aay44938 Artichoke
18	706.5	20.3	735	7 ADM56357	Adm56357 Frog SLC2
19	687.5	19.8	596	7 ABO75066	Abp75066 Pseudomon
20	679.5	19.6	735	7 ADM56276	Adm56276 Mouse SLC
21	679.5	19.6	758	7 ADM56274	Adm56274 Mouse SLC
22	675.5	19.4	311	3 AAY44937	Aay44937 Corn sulp
23	670	19.3	753	7 ADM56359	Adm56359 Pig SLC26
24	668.5	19.2	758	7 ADM56270	Adm56270 Human SLC
25	668	19.2	738	7 ADM56272	Adm56272 Human SLC

26	665.5	19.2	758	3 AAY71067	Aay71067 Human mem
27	664.5	19.1	698	8 ABM84951	Abm84951 Human dia
28	663	19.1	697	8 ABM84952	Abm84952 Human dia
29	662	19.1	751	4 AAG67162	Aag67162 Amino aci
30	662	19.1	4115	5 ABP53570	Abp53570 Human NOV
31	658.5	18.9	746	3 AAY44358	Aay44358 P. chryso
32	657	18.9	721	6 ABG75831	Abg75831 Transport
33	651.5	18.7	328	5 AAE22124	Aae22124 Human 543
34	651.5	18.7	328	6 ADA20211	Ada20211 Sulphate
35	650.5	18.7	744	7 ADM57787	Adm57787 Murine pr
36	650.5	18.7	744	8 AD86098	Ad86098 Murine pr
37	650	18.7	744	7 ADM57785	Adm57785 Gerbil pr
38	650	18.7	744	8 AD86096	Ad86096 Gerbil pr
39	649.5	18.7	716	8 ABM84960	Abm84960 Human dia
40	647.5	18.6	739	2 AAY08929	Aay08929 Rat PrDST
41	647.5	18.6	739	7 ADD46847	Add46847 Rat Prote
42	643.5	18.5	688	8 ABM84953	Abm84953 Human dia
43	642.5	18.5	739	5 AAU74765	Aau74765 Human SLC
44	641.5	18.5	739	2 AAY08928	Aay08928 Human LTD
45	641.5	18.5	739	5 AAU74762	Aau74762 Human SLC

## ALIGNMENTS

## RESULT 1

AAAY44943

ID AAY44943 standard; protein; 680 AA.

XX AAY44943;

XX AC

XX 23-MAY-2000 (first entry)

XX XX

XX Soybean sulphate permease-2.

XX XX

XX Sulphate Permease; sulphate assimilation protein; soybean; probe;

XX mapping; marker; plant breeding; chimeric gene; transgenic plant;

XX antibody; screen.

XX Glycine max.

XX OS

XX WO200004154-A2.

XX PD

XX 27-JAN-2000.

XX PF

XX 13-JUL-1999; 99WO-US015810.

XX PR

XX 14-JUL-1998; 98US-0092833P.

XX XX

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX PA

XX Allen SM, Falco SC, Thorpe CJ;

XX XX

XX WPI; 2000-195025/17.

XX DR

XX N-PSDB; AAY50489.

XX XX

XX Nucleic acid fragments encoding sulfate assimilation proteins in plants

XX PT

XX and seeds useful as probes for isolating cDNAs and genes encoding

XX PT

XX homologous proteins, in producing transgenic plants.

XX PS

XX Claim 6; Page 53-55; 79pp; English.

XX XX

XX The present amino acid sequence is the soybean sulphate permease, a

XX CC

XX sulphate assimilation protein. This sequence is obtained from sfl1

XX CC

XX library, clone sfl1.pk0043.g10, derived from soybean immature flower. It

XX CC

XX has 75% sequence identity to Arabidopsis thaliana sulphate permease (gi

XX CC

XX 2285885). This sequence is used as a probe to isolate other plant

XX CC

XX sulphate assimilation proteins, for genetic and physical mapping of

XX CC

XX related genes and as markers of traits linked to the gene. This is useful

XX CC

XX for plant breeding and to construct chimeric genes, used to create

XX CC

XX transgenic plants with altered levels of sulphate permease. The sulphate

XX CC

XX permease peptides are useful for producing antibodies, that are used to

XX CC

XX screen and isolate cDNA clones

[illegible]

PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144065P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145152P.  
PR 22-JUL-1999; 99US-0145153P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0147192P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146336P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148117P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.

PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 76.0%; Score 2641.5; DB 3; Length 658;  
Best Local Similarity 75.9%; Pred. No. 2e-242;  
Matches 500; Conservative 81; Mismatches 75; Indels 3; Gaps 3;

Qy 24 MGSVDYEPYPLGMNFRVHQ-VEVPPPPFFKSKYSLKETFPDPLRPFQKPKASKKF 82  
Db 1 MGTEDYTFQCAEELHRRHHTVEAPQPFLKSLQYSVKETLFPDPPFQKQNASRK 60  
Qy 83 MLGLQFFPIFEWAPKYYTQFLKADLIAGTTIASLAIPOGISYAKLANLPPILGLYSSEI 142  
Db 61 VLGLKYPLPIFEWAPRYNLKFFKSDLIAGTTIASLAIPOGISYAKLANLPPILGLYSFV 120  
Qy 143 PPLIYAMGSSRDLAGVTAVGSLMGSMLSNAVDNEDPKLYLHLAFTATLFAGVFQAA 202  
Db 121 PPLVAVLGSSRDLAGVTAVASLLTGAMLSKEVDAEKOPKLYLHLAFTATFAGVLEAS 180  
Qy 203 LGLFRLGLIVDFLSHATIIIGFMGAATVVCLOQLKSLILGLEHETHGADIISVMSRVTQT 262  
Db 181 LGIFRLGFIIVDFLSHATIVGFMGAATVVSLOQLKGIIFGLKHFSTDVVISVMSRVSQT 240  
Qy 263 HEWRMESAVLGCVFIIFLLSTRYFSKRPFRFFWVSAMAPLTSVILGSLVYFTHAEKHG 322  
Db 241 HEWRMESVGLGCGFLFLLSTRYFSIKKPKFFWVAAMAPLTSVILGSLVYFTHAEKHG 300  
Qy 323 EVIGELKKGLNPPSLTNLVFVSPYMTTAVKTGIWVGIIISLAEGIAVGRSPAMKYNNIDG 382  
Db 301 QVIGDLKKGLNPLSGSDLIFTSPYMTAVKTGIITGIIAEGVAVGRSPAMKYNNIDG 360  
Qy 383 NKEMIAIGTNVVGSTCYLTTGPFPSAVNNYAGCKTAASNIINSLAVMLLLFLTPL 442  
Db 361 NKEMIAFGMMNIVGSTCYLTTGPFPSAVNNYAGCKTAMSNIVMAIAVMFLLFLTPL 420  
Qy 443 FHYTPLVLSAIIIVSAMLGLIDYEAHILFKVDKDFVVCMSAYIGVVFSGVIGLVI 502

Db 421 FHYTPLVLSAIIISAMGLIDYQAAIHLWKVDKDFLVCMSAYGVGVFSVEIGLVAV 480  
QY 503 VISLRLVLLFIAPRPTVLGNIPNSVIYRNVEHYQNAKHVPGMLLILEIDAPIYFANASYL 562  
Db 481 AISLARLLLVSRPKTAVKGNIPNSMIYRNTEQYPSRTPVPGILILEIDAPIYFANASYL 540  
QY 563 RERITRWIDEEERIKATGETSIQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVL 622  
Db 541 RERIIRWIDEEERVKSGESSIQYIILDSAVGNIDTSGISMMVLEIKKVIDRALKVL 600  
QY 623 VNPVSEVMKLNKSKF-QNHLGKKWIIYLTVEEAVGACNFNLRASKNPNPKDTEGNV 680  
Db 601 SNPKGEVVKLTTRSKFTGDHKGWMLTVGEAVEACSYMLHTFKTEP-ASKNEPNV 658  
  
RESULT 3  
AAG31621  
ID AAG31621 standard; protein; 658 AA.  
XX AC AAG31621;  
XX AC AAG31621;  
DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38004.  
DE DE Protein identification; signal transduction pathway; metabolic pathway;  
KW KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW KW termination sequence.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 03-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137528P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.



PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148694P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148694P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 75.9%; Score 2638.5; DB 3; Length 658;		
Best Local Similarity 75.7%; Pred. No. 3.8e-242;		
Matches 499; Conservative 82; Mismatches 75; Indels 3; Gaps 3;		
Qy	24	MGSVDVEYPLGMNPFERVHQ-VEVPPPPQPFKSLKYSLKETFPDPLROFKNKPASKKF 82
Db	1	MGTEDTFFQGAELHRRHHTVEAPQOPFLKSLQYSVKETLFPDPPFRQFNQASRKF 60
Qy	83	MLGLQFFFFIFEWAPKYTFQFLKADLIAGITIASLAIPQGISYAKLANLPPIILGLSSFI 142
Db	61	VLGLKYFLPIFEWAPRYNLKFFKSDLIAGITIASLAIPQGISYAKLANLPPIILGLSSVF 120
Qy	143	PPLIYAMGSSRDIAVCTVAVGSLLMGSMLSNADPNEDPKLYLHLAFTATLFGVFOAA 202
Db	121	PPLVYAVLGSRRDLAVGTVAVASLLTGAMLSKEVDKOPKLYLHLAFTATTFAGVLEAS 180
Qy	203	LGLFRLGLIVDFLSHATIIIGFMGGAATVVCLOQKSLILGLEHFTHGADIIISVMSVFTQT 262
Db	181	LGIFRLGFIVDFLSHATIIIGFMGGAATLVSLOQKGI FGLKHFTDSTDVSVMSVFSQT 240
Qy	263	HEWRWESAVLGCVFIFFLLLSTRYFSKKRPRFFWVSAMAPLTSVILGSLLYVFTHAERHGV 322
Db	241	HEWRWESVGLGCGFLFLLSTRYFSIKKPKFFWVAAMAPLTSVILGSLLYVFTHAERHGV 300
Qy	323	EVIGELKKGLNPPSLTNLNVSPYMTAVKTIIVGIIISLAEGIAVGRSPAMYKNYNIDG 382
Db	301	QVIGDLKKGLNPLSGSDLIFTSPYMTAVKTIIGITIIAALAEVAVGRSPAMKYNYNIDG 360
Qy	383	NKEMIAIGTMNVVGSFTSCVLTTPFSRAVNNAGCKTAASNIIMSLAVMLTLLFLTPL 442
Db	361	NKEMIAFGMMNIVGSFTSCVLTTPFSRAVNNAGCKTAMSINVMIAVMFTLLFLTPL 420
Qy	443	FHYTPLVLSAIIIVSAMGLIDYEAHILFKVDKDFVVCMSAYIGVFGSVIEGLVIAI 502
Db	421	FHYTPLVLSAIIISAMGLIDYQAAIHLWKVDKDFVLCMSAYVGVFGSVIEGLVAV 480
Qy	503	VISVLRVILFIARPTFVLGNIPNSVIYRNVEHYQNAKHVPGMLIILEIDAPIYFANASYL 562
Db	481	AISIARLLLLFVSRPKTAVKGNIPNSMIYRNTQYPSRTPVPGILILEIDAPIYFANASYL 540
Qy	563	REIRTWIDEEERI KATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVL 622
Db	541	REIRTWIDEEERVVKQSGESSLQYIILDMASVGNIDTSGISMMVEIKKVIDRALKVL 600
Qy	623	VNPVSEVMKKLANKSF-QNHLGKKKIYLYTVEEAVGACNPNLRASKTNPXKDETEGWNV 680
Db	601	SNPKGEVWKLTTRSKFIDGHLGKEMWFLTVGEAVEACSYMLHTFKTEP-ASKNEPNWNV 658
RESULT 4		
AAG31620		
ID	AAG31620 standard; protein; 703 AA.	
XX	AAG31620;	
AC	AAG31620;	
XX	AAG31620;	
DT	17-OCT-2000 (first entry)	
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 38003.	
DE	Arabidopsis thaliana.	
XX	Arabidopsis thaliana.	
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
XX	termination sequence.	
OS	Arabidopsis thaliana.	
XX	Arabidopsis thaliana.	
PN	EP1033405-A2.	
XX	EP1033405-A2.	
PD	06-SEP-2000.	
XX	06-SEP-2000.	
PF	25-FEB-2000; 2000EP-00301439.	
XX	25-FEB-2000; 2000EP-00301439.	
XX	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999; 99US-0123180P.	

PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131149P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0133425P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140891P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143424P.  
PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147933P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148344P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 05-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.

PR	06-OCT-1999;	99US-0157865P.	QY	563	REBITWIDEEERIKATGETSLOYVIDMSAVGNIDTSGISMLEEVKKITERRELOLVL	622
PR	07-OCT-1999;	99US-0158029P.	Db	586	RERIIWIIDEEERVKSGESSLOYIILDMASVGNIDTSGISMMVRIKKYIDRALKLVL	645
PR	08-OCT-1999;	99US-0158232P.				
PR	12-OCT-1999;	99US-0158369P.				
PR	13-OCT-1999;	99US-0159293P.	QY	623	VNPVSEVMKKLANKSF-QNHILGKKIYLVLTVEEAVGACNFNLRASKTNPKKDETEGWNV	680
PR	13-OCT-1999;	99US-0159294P.				
PR	13-OCT-1999;	99US-0159295P.				
PR	14-OCT-1999;	99US-0159329P.	Db	646	SNPKGEVVKLTRSKFIDHILGKHNFLTVGEAVEACSYMLHTFKTEP-ASKNEPWNV	703
PR	14-OCT-1999;	99US-0159330P.				
PR	14-OCT-1999;	99US-0159331P.				
PR	14-OCT-1999;	99US-0159637P.				
PR	14-OCT-1999;	99US-0159638P.				
PR	18-OCT-1999;	99US-0159584P.				
PR	21-OCT-1999;	99US-0160741P.				
PR	21-OCT-1999;	99US-0160787P.				
PR	21-OCT-1999;	99US-0160768P.				
PR	21-OCT-1999;	99US-0160770P.				
PR	21-OCT-1999;	99US-0160814P.				
PR	21-OCT-1999;	99US-0160815P.				
PR	22-OCT-1999;	99US-0160980P.				
PR	22-OCT-1999;	99US-0160981P.				
PR	22-OCT-1999;	99US-0160989P.				
PR	25-OCT-1999;	99US-0161404P.				
PR	25-OCT-1999;	99US-0161405P.				
PR	25-OCT-1999;	99US-0161406P.				
PR	26-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.				
PR	26-OCT-1999;	99US-0161361P.				
PR	28-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161922P.				
PR	28-OCT-1999;	99US-0161932P.				
PR	29-OCT-1999;	99US-0162142P.				
Query Match			75.9%; Score 2638.5; DB 3; Length 703;			
Best Local Similarity			75.7%; Pred. No. 4.2e-242;			
Matches 499; Conservative			82; Mismatches 75; Indels 3; Gaps 3;			
QY	24	MGSDVEYPIGMNFERVHQ-VEVPPQPPFKSLKSKETFPDPPDLRQFKNPKASKKF	82			
Db	46	MGTEDTYFPQGAELHRRHHTVEAPQPFKSLQYSVKETLPDPPDFRQKNQNSRKF	105			
QY	83	MLGLORFPPIFEWAPKYTFQFLKADLIAGITIASLAIPQGISYAKLANLPILGLYSSFI	142			
Db	106	VLGLKYFLPIFEWAPRYNLKFFKSDLIAGITIASLAIPQGISYAKLANLPILGLYSSFV	165			
QY	143	PPLIYAMMGSSRDLAGVTGAVGSLMGMSLSNAVDNEDPKLYLHLAFTATLFAGVFOAA	202			
Db	166	PPLVYAVLGSSRDLAGVTGAVASLLTGAMLSKEVDAEKDPKLYLHLAFTATLFAGVLEAS	225			
QY	203	LGLFRLGLIVDFISHATIIIGFMGGAATVCLQQLKSLGLEHFTHGADIIISVMSRVSFTQT	262			
Db	226	LGIIFRLGFIYDPLSHATIVGFMGGAATLVLSLQQLKGIIFGLKHFTDSTDVIVMSRVSFQT	285			
QY	263	HEWRWESAVLGCYVIFPELLSTRYFSKKRPRFFWVSAMAPLTSVLGSLLVYFTHAEKHGV	322			
Db	286	HEWRWESVLCGCFELFELLSTRIFS IKKAFFWAAWAPLTSVLGSLLVYFTHAEKHGV	345			
QY	323	EVIGELKGLNPPSLTNLNVFSPYMTTAVKTGIWVGIIISLAEGIAVGRSPAMYKNYNIDG	382			
Db	346	QVIGDLKGLNPLSGSDLIFTSPYMTTAVKTGLITGIIALAEVAVGRSPAMFKNYNIDG	405			
QY	383	NKEMIALGTWNVGSPFSCYLTTPGFSRSVAVNTAGCKTAASNIINSLAWMLFLFLTPL	442			
Db	406	NKEMIAFGMNIVGSPFSCYLTTPGFSRSVAVNTAGCKTAMSNIVMAIWMFLFLFLTPL	465			
QY	443	FHTPLVVLSAITVSAMGLIDIEAAHLKPKVDKDFVVCWMSAYIGVFGVSIGIIVIAI	502			
Db	466	FHTPLVVLSAITVSAMGLIDIQAAHLKWKVDKDFVVCWMSAYGVGVFGVSIGLIVAV	525			
QY	503	VISVRLVLLFIARPRTEVLGNINPSVIYRNVHQNAKHVPGMLLILEIDAPIYFANASYL	562			
Db	526	AISIARLLLLFVSRPKTAVKGNIPNSMIYRNTQYPSRSTVPGIILILEIDAPIYFANASYL	595			
RESULT 5						
ID	AA44936					
XX	AA44936 standard; protein; 579 AA.					
AC	AA44936;					
XX	23-MAY-2000 (first entry)					
DE	Corn sulphate permease-2.					
KW	Sulphate Permease; sulphate assimilation protein; corn; contig; probe;					
KW	mapping; marker; plant breeding; chimeric gene; transgenic plant;					
OS	Zea mays.					
XX	WO200004154-A2.					
PD	27-JAN-2000.					
XX	13-JUL-1999; 99WO-US015810.					
XX	14-JUL-1998; 98US-0092833P.					
XX	(DUPO ) DU PONT DE NEMOURS & CO E I.					
PI	Allen SM, Falco SC, Thorpe CJ;					
DR	WPI; 2000-195025/17.					
DR	N-PSDB; AA250482.					
PT	Nucleic acid fragments encoding sulfate assimilation proteins in plants					
PT	and seeds useful as probes for isolating cDNAs and genes encoding					
XX	homologous proteins, in producing transgenic plants.					
PS	Claim 6; Page 42-43; 79pp; English.					
XX	The present amino acid sequence is the corn sulphate permease, a sulphate					
CC	assimilation protein. This sequence is from a contig composed of clones					
CC	crln.pk0015.a2, p0006.cbyrs25rb, p0072.comhc25r, p0091.cmard29r and					
CC	p0092.chwat43r from crln, p0006, p0072, p0091, and p0092 libraries,					
CC	respectively. It has 70% sequence identity to Arabidopsis thaliana					
CC	sulphate permease (gi 2285885). This sequence is used as a probe to					
CC	isolate other plant sulphate assimilation proteins, for genetic and					
CC	physical mapping of related genes and as markers of traits linked to the					
CC	gene. This is useful for plant breeding and to construct chimeric genes,					
CC	used to create transgenic plants with altered levels of sulphate					
CC	permease. The sulphate permease peptides are useful for producing					
XX	antibodies, that are used to screen and isolate cDNA clones					
SQ	Sequence 579 AA;					
Query Match						
Best Local Similarity			62.3%; Score 2164; DB 3; Length 579;			
Matches 403; Conservative			71.0%; Pred. No. 5.7e-197;			
			86; Mismatches 79; Indels 0; Gaps 0;			
QY	105	KADLIAGITIASLAIPQGISYAKLANLPILGLYSSFIPLIYAMMGSSRDLAGVTAVG	164			
Db	3	ESDLIAGITIASLAIPQGISYAKLANLPILGLYSSFVPLIYALMGSSKDLAVGTAVA	62			
QY	165	SLIMGMSLSNAVDNEDPKLYLHLAFTATLFAGVFOALGLFRLGLIVDFLSHATIIIGFM	224			
Db	63	SLLISSMLGSEVSPTEPNVLYLHLAFTATLFAGVFOASGLLRLGLIVDFLSHATIIIGFM	122			

Qy 225 GGAATVVCLOQLKSIIGLEHFTGADIISVMRSVFTQTHEWRWESAVLGCVFIFLLSTR 284  
Db 123 AGAATVVCLOQLKGLVHFTTSTDVSVMESVFSQTHQWRWESVLLGCGFLFLVTR 182  
Qy 285 YFSKKRPRPFWSAMAPLTSVIIIGSLVYTHAEKGVVEIGELKGLNPPSLTNLVFS 344  
Db 183 FISKRRPKLFWISAAAPLTSWSILGSLVYLTHAENHGVIGYLRKGLNPPSVTSLOFSP 242  
Qy 345 PYMTAVKTIQVVGITISLAGIAGVSGSFAMKYKNYNIDGNKEMIAIGTMNVVGSFTSCYLT 404  
Db 243 PYMMLAKTGIITGVIALAGIAGVSGSFAMKYNHMTDNKEMIAIGTMNVVGLSLTSCYLT 302  
Qy 405 TGPFSRNVNAGCKTAASNIIMSLAVMLTLFLTPHFTPLVVLVLSAIIIVSAMGLID 464  
Db 303 TGPFSRNVNAGCTAMSNVMSLAVMVTLLFLTPHFTPLVVLVLSAIIIVSAMGLVD 362  
Qy 465 YEAAIHLFKVDKDFVVCMSAYIGVFGSVEIGLVAIVLSURVLLFIARPTTFVIGMI 524  
Db 363 FGAALHLWRVDKDFVCAGAYLGVVFGSVEGLVAVAVSLRVLFFVARPTTVLIGNI 422  
Qy 525 PMSVIVRNVEHYQNAKHVPGMLILEIDAPIYFANASVLRERITRWIDEEERIKATGETS 584  
Db 423 PGTWVRMDQYAAQVPGVLRVDAPVYFANASYLRERISRWIDDEERTKSGEMG 482  
Qy 585 LQYIIDMGAIGNIDTSGISMLSEVKKITERRELQVLVNPVSEVMKKLNKSKFQNHGK 644  
Db 483 VRYVLDWGAIGSIDTSGTSMDELNKSIDRRGMQIVLANPGSEIMKKLDSKVLQIGH 542  
Qy 645 KWIYLTVEEAVGACNPNLRASKTNPKD 672  
Db 543 EWVFPTVEAVASCDYVLHSHKFGMAKD 570

RESULT 6  
AAG32201  
ID AAG32201 standard; protein; 510 AA.  
XX  
AC AAG32201;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38801.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 23-JUL-1999; 99US-0145276P.  
PR 26-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147933P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150586P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 12-OCT-1999; 99US-0158232P.  
PR 13-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159594P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.

PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0161404P.  
PR 23-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161320P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 59.6%; Score 2070; DB 3; Length 510;  
Best Local Similarity 76.7%; Pred. No. 4.4e-188;  
Matches 392; Conservative 64; Mismatches 53; Indels 2; Gaps 2;  
  
QY 171 MLSNAVDPNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLSHATIIQFMGGAATV 230  
DB 1 MLSKEVDAEKDPKLYLHLAFTATFFAGVLEASLGIFRLGFIIVDFLSHATIVGFMGGAATV 60  
  
QY 231 VCLQQLKSIILGLEBHTHGADIIISVMSRVFTQTHWRWESAVLGCVFIFFLSLSTRYFSKKR 290  
DB 61 VSLQQLKGIPLKXHTDSTDVISVMSRVFSQTHWRWESVGLGCGPFLFLLSTRYFSIKK 120  
  
QY 291 PRFFWVSAMAPLTSVLGSLIVYFTHAEKHGVEVIGELKKGLNPPSLTNLVFVSPYMTTA 350  
DB 121 PKPFVWAAMAPLTSVLGSLIVYFTHAEKHGVEVIGELKKGLNPPSLTNLVFVSPYMTTA 180  
  
QY 351 VKTGIVVGIISLAEGIAVGRSFAMYNKYNIDGNKEMIAICTMNVVGSFTSCYLTGTGPFSSR 410  
DB 181 VKTGLITGIIALAEVAVGRSFAMYNKYNIDGNKEMIAICTMNVVGSFTSCYLTGTGPFSSR 240  
  
QY 411 SAVYNAGCKTAASNIIMSLAVMLTLLFLPLFHYTPLVVLVAIIISAMGLIDYEAH 470  
DB 241 SAVYNAGCKTAMSNIVMAJAVMFTLLFLPLFHYTPLVVLVAIIISAMGLIDYEAH 300  
  
QY 471 LFKVDKEDFVVCMSAYIGUVFGSVEIGLVAIIVISLVLFIARPTFVLGNIPNSVIY 530  
DB 301 LFKVDKEDFVVCMSAYIGUVFGSVEIGLVAIIVISLVLFIARPTFVLGNIPNSVIY 360  
  
QY 531 RNVEHYONAKHVPGLMILEIDAPIYFANASYLRERITRWIDEERERIKATGETSLOQVII 590  
DB 361 RNTEQYPSRTVEGIIILEIDAPIYFANASYLRERITRWIDEERERIKATGETSLOQVII 420  
  
QY 591 DMSAVGNIDTSGISMLEEVKKITERRELBQLVLVNPVSEVMKKLNKSKF-QNHILGKKWIYL 649  
DB 421 DMSAVGNIDTSGISMLEEVKKITERRELBQLVLVNPVSEVMKKLNKSKF-QNHILGKKWIYL 480  
  
QY 650 TVEEAVGACNPNLRASKTNPKKDETEGWNV 680  
DB 481 TVGEAVEACSYMLHTFKTEP-ASKNEPWNV 510  
  
RESULT 7  
AAG31622  
ID AAG31622 standard; protein; 510 AA.  
XX  
AC AAG31622;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38005.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.

XX EP1033405-A2.  
PN  
XX  
XX  
PD  
XX  
PF 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 03-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130049P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 14-JUN-1999; 99US-0138847P.  
PR 16-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.



```
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 08-OCT-1999; 99US-0157753P.
PR 08-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 13-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 59.5%; Score 2067; DB 3; Length 510;
Best Local Similarity 76.5%; Pred. No. 8.5e-188;
Matches 391; Conservative 65; Mismatches 53; Indels 2; Gaps 2;

QY 171 MLSNADVPNEDPKLYLHLAFTATLFAGVFOAALGLFRGLIVDFLGHATTIIGFMGGAATV 230
DB 1 MLSKEVDAEKDPKLYLHLAFTATLFAGVLEASLGIFRLGFIIVDFLGHATTIIGFMGGAATL 60

QY 231 VCLQQLKSILGLEBHTHGDADIIISVMSRVFTQTHEWRMESAVLGCVFIFLLSTRYFSKRR 290
DB 61 VSLQQLKGFGLKHTDSTDVISMVSFVSQTHEWRMESGVLCGFLFLLSTRYFSIKK 120

QY 291 PRFFWVSAMAPLTSVILGSLVLYFTAEKHGVVIGELKGLNPPSLTNLVFVSPWMTA 350
DB 121 PKFFVAAAMAPLTSVILGSLVLYFTAEKHGVVIGELKGLNPLSGSLDIFTPTMSTA 180

QY 351 VKTGIVGVGIIISLAEGTAGVGRFAMYNKYNINDGNKEMIAITGMNVGSGFTSCYLLTGPFSR 410
DB 181 VKTGLITGIIAELGAVAGVGRFAMYNKYNINDGNKEMIAFGMMNIVGSGFTSCYLLTGPFSR 240

QY 411 SAVYNAGCKTAASNIIMSLAVMLTLFTPLPHYTPLVLSAIIIVSAMGLGIDYEAH 470
DB 241 SAVYNAGCKTAMSNIVMAVMTLTLFTPLPHYTPLVLSAIIISAMGLGIDYQAAH 300

QY 471 LFKVDKDFVCMVSAYIGVVGVSVEIGLVIAIVISVLRLFTARPRFTVLGNIPNSVIY 530
DB 301 LMKVDKDFVCMVSAYIGVVGVSVEIGLVVAIVASIAARLLFVSRPKTAVKGNIPNSMIY 360

QY 531 RNVEHYONAKHVPGLILEIDAPIYANASYLPERITRITDEEERIKATGETSLOYVII 590
DB 361 RNTEQVPSSTVPGLILEIDAPIYANASYLPERIIRWIDEEERVKQSGESSLOYIIL 420

QY 591 DMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNSKFP-QNHILGKKWIYL 649
DB 421 DMSAVGNIDTSGISMVVEIKKVIDRRALKVLVLSNPKGEVVKVKKLTRSKFTGDHGLKEMFL 480

QY 650 TYEEAVGACNFNLASKTNPKKDETEGMNV 680
DB 481 TVGEAVEACSYMLHTFKTEP-ASKNEPMNV 510

RESULT 8
AAY44942
ID AAY44942 standard; protein; 621 AA.
AC AAY44942;
XX
XX
DT 23-MAY-2000 (first entry)
XX
DE Soybean sulphate permease-1.
KW Sulphate Permease; sulphate assimilation protein; soybean; probe;
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;
KW antibody; screen.
XX
OS Glycine max.
XX
PN WO200004154-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015810.
PR 14-JUL-1998; 98US-0092833P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
PA Allen SM, Falco SC, Thorpe CJ;
PI
XX
DR WPI; 2000-195025/17.
DR N-PSDB; AA250488.
XX
XX
PT Nucleic acid fragments encoding sulfate assimilation proteins in plants
PT and seeds useful as probes for isolating cDNAs and genes encoding
PT homologous proteins, in producing transgenic plants.
XX
XX Claim 6; Page 50-52; 79pp; English.
XX
CC The present amino acid sequence is the soybean sulphate permease, a
CC sulphate assimilation protein. This sequence is obtained from ses2w
CC library, clone ses2w.pk0031.b3, derived from soybean embryogenic
CC suspension 2 weeks after subculture. It has 61% sequence identity to
CC Arabidopsis thaliana sulphate permease (gi 4579913). This sequence is
CC used as a probe to isolate other plant sulphate assimilation proteins,
CC for genetic and physical mapping of related genes and as markers of
CC traits linked to the gene. This is useful for plant breeding and to
CC construct chimeric genes, used to create transgenic plants with altered
CC levels of sulphate permease. The sulphate permease peptides are useful
CC for producing antibodies, that are used to screen and isolate cDNA clones
XX
XX Sequence 621 AA;

Query Match 54.7%; Score 1901; DB 3; Length 621;
Best Local Similarity 59.0%; Pred. No. 7.7e-172;
Matches 359; Conservative 106; Mismatches 143; Indels 0; Gaps 0;

QY 48 PPQPFKSLKYSKETFPPDDPLRQFNKPKASKMKMLGLOFPFPFEWAPKTYFQFLKAD 107
DB 3 PHQTTLHKLRHVSEIFFPDDPLRQFNKPKASKMKMLGLOFPFPFEWAPKTYFQFLKAD 62

QY 108 LIAGTTIASLAIPQGISYAKLANLPILGLYSSTPPLIYAMGSSRDIAVGVAGSLL 167
DB 63 LISGLTIIASLAIPQGISYAKLANLPILGLYSSTPPLIYAMGSSRDIAVGVAGSLL 122
```

Qy	168	MGSMLSNAVDNEDPKLYLHLAFATLIFAGVQQAALGFLRGLIVDFLSHATIIIGFWGGA	227	PR	28-APR-1999;	99US-013114499;
Db	123	MGSMLSKISVTPQEPILYGLAPATATFAFAGVQQAASLGILRGLFVIDFLSKATIVGFTGGA	182	PR	30-APR-1999;	99US-01320408P;
Qy	228	ATVVCLOOLKSILGLEHFTHGADIISVMRSVFTCTHWRWESAVLGCVFIFLLSTRYFS	287	PR	04-MAY-1999;	99US-01324844P;
Db	183	AIIVSLOOLKGLGIVHFTSKMQIIPVTISVFKQRHEWSQTIILGLSGFLVFLTLTRHIS	242	PR	05-MAY-1999;	99US-0132485P;
Qy	288	KCRPRFPWWSAMAPLTSVILGSLVYFTHAKHGVEVIGELKGLNPPSLTNLVFVSPYM	347	PR	06-MAY-1999;	99US-0132486P;
Db	243	LRPKLFWWSAAPLTSVILSTILVFLLRNKTQIISVIGHLPKGVNPPSANMLYFNGPYL	302	PR	07-MAY-1999;	99US-0132863P;
Qy	348	TTAVKGTIVVGIISLAEGIAVGRSFAMKYKNYNDGNKEMIAIGTMNVVGSFTSCYLTITGP	407	PR	11-MAY-1999;	99US-0134256P;
Db	303	GLAIKGTGIITGILSLTSGIAVGRFTFASLKNYQVDGNKEMWAIGLMNIAGSCSCYYVTGS	362	PR	14-MAY-1999;	99US-0134218P;
Qy	408	FSRSVNNYAGCKTAASNIIIMSLAVMLTLLFLPFLHYTFLVNLVLSAIIIVSAMGLIDYEA	467	PR	14-MAY-1999;	99US-0134221P;
Db	363	FSRSVNNYAGAOQTVSNIIIMAAVLTLLFLMPLFYFTPNVNLAAIIITAVIGLIDYQS	422	PR	14-MAY-1999;	99US-0134221P;
Qy	468	AJHLFKVDKDFVVCMSAYIGVWFGSVEIGIAVIAIVSLRVLVLTARPTFVLGNIPNS	527	PR	18-MAY-1999;	99US-0134370P;
Db	423	AYLKWVKDKLDFLACLCSFFGVLFISVPLGLGIAVIVSLIKILLHTRPNTLVGLNPGT	482	PR	18-MAY-1999;	99US-0134341P;
Qy	528	VIYRNVHYQNAKHVPGMLILEIDAPIYFANASYLRERIRRWIDEBEERIKATGETSLOY	587	PR	20-MAY-1999;	99US-0135124P;
Db	483	QIFHNINQYKKALRVPSFLILAVESPIYFANSYIQLERILRWVREBEHIKANNAPLKC	542	PR	21-MAY-1999;	99US-0135353P;
Qy	588	VIIDMSAVGNIDTSGISMLSEVKKITERRELQVLVNPVSEVMKKLKSFKQHLGKKWI	647	PR	24-MAY-1999;	99US-0135629P;
Db	543	IILDMTAVTATDTSGLDTLCELRMLEKRSLEFLVLANPVGNVMEKHLKHSNILDSPGLKGV	602	PR	25-MAY-1999;	99US-0136021P;
Qy	648	YLTVEEAV 655		PR	27-MAY-1999;	99US-0136392P;
Db	603	YLTVEEAV 610		PR	28-MAY-1999;	99US-0136782P;
RESULT 9				PR	01-JUN-1999;	99US-0137222P;
AAG29247				PR	03-JUN-1999;	99US-0137502P;
ID	AAG29247 standard; protein; 631 AA.			PR	04-JUN-1999;	99US-0137724P;
AC	AAG29247;			PR	07-JUN-1999;	99US-0137502P;
XX	DT			PR	08-JUN-1999;	99US-0138094P;
XX	DT			PR	10-JUN-1999;	99US-0138847P;
XX	DE			PR	14-JUN-1999;	99US-0139119P;
XX	DE			PR	16-JUN-1999;	99US-0139453P;
XX	DE			PR	17-JUN-1999;	99US-0139492P;
XX	DE			PR	18-JUN-1999;	99US-0139454P;
XX	DE			PR	18-JUN-1999;	99US-0139455P;
XX	DE			PR	18-JUN-1999;	99US-0139456P;
XX	DE			PR	18-JUN-1999;	99US-0139457P;
XX	DE			PR	18-JUN-1999;	99US-0139458P;
XX	DE			PR	18-JUN-1999;	99US-0139459P;
XX	DE			PR	18-JUN-1999;	99US-0139460P;
XX	DE			PR	18-JUN-1999;	99US-0139461P;
XX	DE			PR	18-JUN-1999;	99US-0139462P;
XX	DE			PR	18-JUN-1999;	99US-0139463P;
XX	DE			PR	18-JUN-1999;	99US-0139750P;
XX	DE			PR	18-JUN-1999;	99US-0139763P;
XX	DE			PR	21-JUN-1999;	99US-0139817P;
XX	DE			PR	22-JUN-1999;	99US-0139899P;
XX	DE			PR	23-JUN-1999;	99US-0140353P;
XX	DE			PR	23-JUN-1999;	99US-0140354P;
XX	DE			PR	24-JUN-1999;	99US-0140695P;
XX	DE			PR	28-JUN-1999;	99US-0140823P;
XX	DE			PR	29-JUN-1999;	99US-0140991P;
XX	DE			PR	30-JUN-1999;	99US-014287P;
XX	DE			PR	01-JUL-1999;	99US-0141842P;
XX	DE			PR	01-JUL-1999;	99US-0142154P;
XX	DE			PR	02-JUL-1999;	99US-0142055P;
XX	DE			PR	06-JUL-1999;	99US-0142390P;
XX	DE			PR	08-JUL-1999;	99US-0142803P;
XX	DE			PR	09-JUL-1999;	99US-0142920P;
XX	DE			PR	12-JUL-1999;	99US



XX AAG32202; 18-JUN-1999; 99US-0139462P.  
AC 18-JUN-1999; 99US-0139463P.  
XX 18-JUN-1999; 99US-0139750P.  
DT 18-JUN-1999; 99US-0139763P.  
XX 21-JUN-1999; 99US-0139817P.  
DE 22-JUN-1999; 99US-0139899P.  
XX 23-JUN-1999; 99US-0140353P.  
KW 23-JUN-1999; 99US-0140354P.  
KW 24-JUN-1999; 99US-0140695P.  
KW 24-JUN-1999; 99US-0140823P.  
XX 28-JUN-1999; 99US-0140823P.  
OS 29-JUN-1999; 99US-0140991P.  
XX 30-JUN-1999; 99US-0141287P.  
PN 01-JUL-1999; 99US-0141842P.  
XX 01-JUL-1999; 99US-0142154P.  
XX 02-JUL-1999; 99US-0142055P.  
XX 06-JUL-1999; 99US-0142390P.  
XX 08-JUL-1999; 99US-0142803P.  
XX 09-JUL-1999; 99US-0142920P.  
PF 12-JUL-1999; 99US-0142977P.  
XX 13-JUL-1999; 99US-0143542P.  
XX 14-JUL-1999; 99US-0143624P.  
XX 15-JUL-1999; 99US-0144005P.  
XX 16-JUL-1999; 99US-0144085P.  
XX 16-JUL-1999; 99US-0144086P.  
XX 19-JUL-1999; 99US-0144325P.  
XX 19-JUL-1999; 99US-0144331P.  
XX 19-JUL-1999; 99US-0144332P.  
XX 19-JUL-1999; 99US-0144333P.  
XX 19-JUL-1999; 99US-0144334P.  
XX 19-JUL-1999; 99US-0144335P.  
XX 20-JUL-1999; 99US-0144352P.  
XX 20-JUL-1999; 99US-0144632P.  
XX 20-JUL-1999; 99US-0144884P.  
XX 21-JUL-1999; 99US-0144814P.  
XX 21-JUL-1999; 99US-0145086P.  
XX 21-JUL-1999; 99US-0145088P.  
XX 22-JUL-1999; 99US-0145085P.  
XX 22-JUL-1999; 99US-0145087P.  
XX 22-JUL-1999; 99US-0145089P.  
XX 22-JUL-1999; 99US-0145192P.  
XX 23-JUL-1999; 99US-0145145P.  
XX 23-JUL-1999; 99US-0145218P.  
XX 23-JUL-1999; 99US-0145224P.  
XX 26-JUL-1999; 99US-0145276P.  
XX 27-JUL-1999; 99US-0145913P.  
XX 27-JUL-1999; 99US-0145918P.  
XX 27-JUL-1999; 99US-0145919P.  
XX 28-JUL-1999; 99US-0145951P.  
XX 02-AUG-1999; 99US-0146386P.  
XX 02-AUG-1999; 99US-0146388P.  
XX 02-AUG-1999; 99US-0146389P.  
XX 03-AUG-1999; 99US-0147038P.  
XX 04-AUG-1999; 99US-0147204P.  
XX 04-AUG-1999; 99US-0147302P.  
XX 05-AUG-1999; 99US-0147192P.  
XX 05-AUG-1999; 99US-0147260P.  
XX 06-AUG-1999; 99US-0147303P.  
XX 06-AUG-1999; 99US-0147416P.  
XX 09-AUG-1999; 99US-0147493P.  
XX 09-AUG-1999; 99US-0147935P.  
XX 10-AUG-1999; 99US-0148171P.  
XX 11-AUG-1999; 99US-0148319P.  
XX 12-AUG-1999; 99US-0148341P.  
XX 13-AUG-1999; 99US-0148565P.  
XX 13-AUG-1999; 99US-0148684P.  
XX 16-AUG-1999; 99US-0149368P.  
XX 17-AUG-1999; 99US-0149175P.  
XX 18-AUG-1999; 99US-0149426P.  
XX 20-AUG-1999; 99US-0149722P.  
XX 20-AUG-1999; 99US-0149723P.  
XX 20-AUG-1999; 99US-0149923P.  
XX 23-AUG-1999; 99US-0149902P.  
XX 23-AUG-1999; 99US-0149930P.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 38802.  
Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 23-APR-1999; 99US-0130891P.  
XX 28-APR-1999; 99US-0131449P.  
XX 30-APR-1999; 99US-0132048P.  
XX 30-APR-1999; 99US-0132407P.  
XX 04-MAY-1999; 99US-0132484P.  
XX 05-MAY-1999; 99US-0132485P.  
XX 06-MAY-1999; 99US-0132486P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132863P.  
XX 11-MAY-1999; 99US-0134256P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134768P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136392P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137222P.  
XX 03-JUN-1999; 99US-0137228P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138847P.  
XX 14-JUN-1999; 99US-0139119P.  
XX 16-JUN-1999; 99US-0139452P.  
XX 16-JUN-1999; 99US-0139453P.  
XX 17-JUN-1999; 99US-0139453P.  
XX 18-JUN-1999; 99US-0139454P.  
XX 18-JUN-1999; 99US-0139455P.  
XX 18-JUN-1999; 99US-0139456P.  
XX 18-JUN-1999; 99US-0139457P.  
XX 18-JUN-1999; 99US-0139458P.  
XX 18-JUN-1999; 99US-0139459P.  
XX 18-JUN-1999; 99US-0139460P.  
XX 18-JUN-1999; 99US-0139461P.

Qy	464	DYEAATHLKFVKDFVVCMSAYIGVFGSVEIGLVIAIVISVLRLVLLFIARPTFVLGN	523
Db	241	DYQAAIHLWKVKDFVVCMSAYIGVFGSVEIGLVIAIVISVLRLVLLFIARPTFVLGN	300
Qy	524	IPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLRETRTWIDEEERIKATGET	583
Db	301	IPNSMIYRNTEQYPSRTPVGLILEIDAPIYFANASYLRETRTWIDEEERIKATGET	360
Qy	584	SLQYIILDSAVGNIDTSGISMLBEEVKKITERRELQVLVNPVSEVMKLNKSKF-QNHIL	642
Db	361	SLQYIILDSAVGNIDTSGISMLBEEVKKITERRELQVLVNPVSEVMKLNKSKF-QNHIL	420
Qy	643	GKKWIYLTVEEAVGACNFNLRSKTNPKKDETEGMNV	680
Db	421	GKWMELTVGEAVEACSYMLHTFKTEP-ASKNEPWNV	457
RESULT 11			
AA44935			
ID	AA44935	standard; protein; 688 AA.	
XX	AA44935;		
AC	AA44935;		
XX	23-MAY-2000	(first entry)	
XX	Corn sulphate permease-1.		
XX	Sulphate Permease; sulphate assimilation protein; corn; contig; probe;		
KW	mapping; marker; plant breeding; chimeric gene; transgenic plant;		
KW	antibody; screen.		
XX	Zea mays.		
XX	WO200004154-A2.		
XX	27-JAN-2000.		
XX	13-JUL-1999;	99WO-US015810.	
XX	14-JUL-1998;	98US-0092833P.	
XX	(DUPO ) DU PONT DE NEMOURS & CO E I.		
XX	Allen SM, Falco SC, Thorpe CJ;		
XX	WPI; 2000-195025/17.		
XX	N-PSDB; AA250481.		
PT	Nucleic acid fragments encoding sulfate assimilation proteins in plants		
PT	and seeds useful as probes for isolating cDNAs and genes encoding		
PT	homologous proteins, in producing transgenic plants.		
XX	Claim 6; Page 39-41; 79pp; English.		
XX	The present amino acid sequence is the corn sulphate permease, a sulphate		
CC	assimilation protein. This sequence is from a contig composed of clones		
CC	cbln10.pk0062.b10, ccoln.pk081.h21, ccoln.pk092.12, cscic.pk005.j3,		
CC	p0004.cblej58r, p0089.csdch19r, p0094.csesg12r, p0121.cfrmx30r and		
CC	p0128.cbicj09r from cbln10, ccoln, cscic, p0004, p0089, p0094, p0121 and		
CC	p0128 libraries, respectively. It has 80.2% sequence identity to		
CC	Sporobolus stapifianus sulphate permease(gi 1907270). This sequence is		
CC	used as a probe to isolate other plant sulphate assimilation proteins,		
CC	for genetic and physical mapping of related genes and as markers of		
CC	traits linked to the gene. This is useful for plant breeding and to		
CC	construct chimeric genes, used to create transgenic plants with altered		
CC	levels of sulphate permease. The sulphate permease peptides are useful		
CC	for producing antibodies, that are used to screen and isolate cDNA clones		
XX	Sequence 688 AA;		
Query Match 53.1%; Score 1843.5; DB 3; Length 688;			
Best Local Similarity 56.2%; Pred. No. 2.7e-166;			

Matches	354; Conservative	115; Mismatches	156; Indels	5; Gaps	3;
Qy	41	VHVEVPPQPFKSLKYSKLTETFPDPLRQFNKPKSKKFMGLQGFPPFIPEWAPKYT	100		
Pt	11				
Pt	12				
Pt	13				
Pt	14				
Pt	15				
Pt	16				
Pt	17				
Pt	18				
Pt	19				
Pt	20				
Pt	21				
Pt	22				
Pt	23				
Pt	24				
Pt	25				
Pt	26				
Pt	27				
Pt	28				
Pt	29				
Pt	30				
Pt	31				
Pt	32				
Pt	33				
Pt	34				
Pt	35				
Pt	36				
Pt	37				
Pt	38				
Pt	39				
Pt	40				
Pt	41				
Pt	42				
Pt	43				
Pt	44				
Pt	45				
Pt	46				
Pt	47				
Pt	48				
Pt	49				
Pt	50				
Pt	51				
Pt	52				
Pt	53				
Pt	54				
Pt	55				
Pt	56				
Pt	57				
Pt	58				
Pt	59				
Pt	60				
Pt	61				
Pt	62				
Pt	63				
Pt	64				
Pt	65				
Pt	66				
Pt	67				
Pt	68				
Pt	69				
Pt	70				
Pt	71				
Pt	72				
Pt	73				
Pt	74				
Pt	75				
Pt	76				
Pt	77				
Pt	78				
Pt	79				
Pt	80				
Pt	81				
Pt	82				
Pt	83				
Pt	84				
Pt	85				
Pt	86				
Pt	87				
Pt	88				
Pt	89				
Pt	90				
Pt	91				
Pt	92				
Pt	93				
Pt	94				
Pt	95				
Pt	96				
Pt	97				
Pt	98				

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
PA  
XX  
DR WPI; 2002-718707/78.





PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 14-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140895P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 11-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 18-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 12-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.

[illegible]

PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144326P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145313P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 08-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.

PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 36.6%; Score 1271.5; DB 3; Length 446;  
Best Local Similarity 54.7%; Pred.No.5.2e-112;  
Matches 238; Conservative 93; Mismatches 103; Indels 1; Gaps 1;

QY 224 MGAATVVCLOQLKSLGLEHFTHGADIISVNRSVFTQTHWRWESAVLGCVFIFFLST 283  
DB 1 MGAATVVCLOQLKSLGLEHFTHGADIISVNRSVFTQTHWRWESAVLGCVFIFFLST 60  
QY 284 RYFSKKRPRFPWVSAMAPLTSVILGSLLYVFFTHAEKHGVEVIGELKGLNPPSLTNLVFV 343  
DB 61 RHLMSKKPKLFWVSAGAPLLSVIVSTLLVFFVRAERHGISVIGKLEGLNPPSWNNLQFH 120  
QY 344 SPYMTTAVKTGVVGIISLAEGIAVGRSPAMVKNYNIDGNKEMIAIGTMNVVGSFTSCYL 403  
DB 121 GSHLALVAKTGLVTGIVSLTEGIAVGRTPAALKNVAVDGNKEMIAIGTMNVVGSATSCYV 180  
QY 404 TTGPFSSAVNYNAGCKTAASNIIMSLAVMLTLLFLTPLFHYTPLVLSAIVSAMGLI 463  
DB 181 TTGAFSSAVNNGAKTAVSNIVMSVTVMVTLFLMPLFEVTPNVVLGAIIVTAVIGLI 240  
QY 464 DYEAAIHLEKVDKDFVVCMSAYIGVVGSGVEIGLVIAIVISVLRVLLFIARPTFVLGN 523  
DB 241 DLPAACHWKIDKDFDLVNLCAFFGVIFLSVQNGLAIAVGLSLFKILMQVTRPKWVIMGN 300  
QY 524 IPNSVIYRNVEHYONAKHVPGLILEIDAPIYFANASYLREITRIMIDEEERIKATGET 583  
DB 301 IEGTDIYRDLHHYKQAQRIPIGFLVLSIESPVNPAFNSNYLTERTSRWIECEBEEAEKHS 360  
QY 584 SLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKGLNKSFKQNH-L 642  
DB 361 SLQFLILEMSAVSGVDNNGVSFFKELKKTAKKDIELVFNPLSEVVEKQLQRADEKQFM 420  
QY 643 GKWTYLTVEEAVGA 657  
DB 421 RPEFLFTVAEAVAS 435

Search completed: August 31, 2005, 09:41:28  
Job time : 128 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: August 31, 2005, 08:53:14 ; Search time 33 Seconds  
(without alignments)  
1538.223 Million cell updates/sec

Title: US-10-762-049-18  
Perfect score: 3475  
Sequence: 1 HELARTLSYITHICLLRNTI.....NLRASKINPKKDETEGNNV 680

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3475	100.0	680	4	US-09-720-317A-18
2	2642.5	76.0	658	4	US-09-720-317A-24
3	2575.5	74.1	646	4	US-09-720-317A-26
4	2314	66.6	646	4	US-09-720-317A-28
5	2164	62.3	579	4	US-09-720-317A-4
6	1901	54.7	621	4	US-09-720-317A-16
7	1890.5	54.4	631	4	US-09-720-317A-29
8	1843.5	53.1	688	4	US-09-720-317A-2
9	1798.5	51.8	656	4	US-09-720-317A-20
10	1760.5	50.7	660	4	US-09-720-317A-30
11	1723.5	49.6	660	4	US-09-720-317A-23
12	1503	43.3	644	4	US-09-720-317A-25
13	997	28.7	685	4	US-09-720-317A-31
14	961.5	27.7	593	4	US-09-720-317A-22
15	731.5	21.1	228	4	US-09-720-317A-8
16	687.5	19.8	233	4	US-09-720-317A-27
17	687.5	19.8	595	4	US-09-252-991A-23812
18	675.5	19.4	311	4	US-09-720-317A-6
19	650.5	18.7	744	4	US-09-785-381-3
20	650	18.7	744	4	US-09-785-381-1
21	643.5	18.5	590	4	US-09-902-540-14944
22	641.5	18.5	748	4	US-09-949-016-10387
23	612	17.6	565	4	US-09-602-787A-616
24	587.5	16.9	780	4	US-09-785-381-11
25	586.5	16.9	764	1	US-08-424-567-2
26	586.5	16.9	764	2	US-08-711-928-2
27	586.5	16.9	764	3	US-09-184-937-2

28	586.5	16.9	790	4	US-09-949-016-11220	Sequence 11220, A
29	576.5	16.6	803	4	US-09-949-016-11498	Sequence 11498, A
30	567	16.3	828	4	US-09-248-796A-20746	Sequence 20746, A
31	549.5	15.8	663	4	US-09-875-811-6	Sequence 6, Appli
32	549	15.8	656	4	US-09-875-811-10	Sequence 10, Appli
33	546.5	15.7	679	4	US-09-875-811-2	Sequence 2, Appli
34	489	14.1	616	4	US-09-543-681A-4421	Sequence 4421, Ap
35	485.5	14.0	605	4	US-09-875-811-8	Sequence 8, Appli
36	485	14.0	598	4	US-09-875-811-12	Sequence 12, Appli
37	482.5	13.9	621	4	US-09-875-811-4	Sequence 4, Appli
38	454	13.1	970	4	US-09-795-927-7	Sequence 7, Appli
39	449	12.9	535	4	US-09-252-991A-21805	Sequence 21805, A
40	426.5	12.3	472	4	US-09-902-540-13374	Sequence 13374, A
41	402.5	11.6	575	4	US-09-438-185A-1015	Sequence 1015, Ap
42	381	11.0	159	4	US-09-720-317A-14	Sequence 14, Appli
43	369	10.6	566	4	US-09-543-681A-4544	Sequence 4544, Ap
44	363	10.4	533	4	US-09-107-532A-5244	Sequence 5244, Ap
45	360.5	10.4	147	4	US-09-720-317A-10	Sequence 10, Appli

## ALIGNMENTS

RESULT 1  
US-09-720-317A-18  
; Sequence 18, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 18  
; LENGTH: 680  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-720-317A-18

Query Match		100.0%;	Score 3475;	DB 4;	Length 680;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 680;		Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	HELARTLSYITHICLLRNTI	I	EDMGSDVDEYEP	PLGMNFRVHQVEVPPQPPFKSLKYSL 60
Db	1	HELARTLSYITHICLLRNTI	I	EDMGSDVDEYEP	PLGMNFRVHQVEVPPQPPFKSLKYSL 60
Qy	61	KETFFPDDPLROFKNPKASKKMLGQFFPF	F	EWAPKTYTQFLKADLAGITIASLAIP 120	
Db	61	KETFFPDDPLROFKNPKASKKMLGQFFPF	F	EWAPKTYTQFLKADLAGITIASLAIP 120	
Qy	121	QGISYAKLANLPILGLYSFIPPLIYAMGSSRD	L	AVGTAVGSLMGSMLSNAVDPNE 180	
Db	121	QGISYAKLANLPILGLYSFIPPLIYAMGSSRD	L	AVGTAVGSLMGSMLSNAVDPNE 180	
Qy	181	DPKLYLHLAFTATLFAGVFOAALGLFRLGLI	V	DFLSHATIIIGPMGAATVVCLOQLKSIL 240	
Db	181	DPKLYLHLAFTATLFAGVFOAALGLFRLGLI	V	DFLSHATIIIGPMGAATVVCLOQLKSIL 240	
Qy	241	GLEHFTHGADIISVMRSVFTQHEWRWESAVL	G	CVFIFFLLSTRYFSKRPFRFWWSAMA 300	
Db	241	GLEHFTHGADIISVMRSVFTQHEWRWESAVL	G	CVFIFFLLSTRYFSKRPFRFWWSAMA 300	
Qy	301	PLTSTVLGSLLVYFTHAEKHGVEVIGELKKGL	N	PPSLTNLNVFVSPYMTTAVKTGIUVGII 360	
Db	301	PLTSTVLGSLLVYFTHAEKHGVEVIGELKKGL	N	PPSLTNLNVFVSPYMTTAVKTGIUVGII 360	

Qy	361	SLAEGIAVGRSFAMYYKQNNINDGNKEMIAIGTMNVVGSFTSCYLTGTPFSRSAVNNAGCK	421	421	TAASNIIMSLAVNMLTLLFLTPLFPHYTPPLAVLSAIIIVSAMIGLIDYEAAIHLFKVDKDFEV	480
Db	361	SLAEGIAVGRSFAMYYKQNNINDGNKEMIAIGTMNVVGSFTSCYLTGTPFSRSAVNNAGCK	420	421	TAASNIIMSLAVNMLTLLFLTPLFPHYTPPLAVLSAIIIVSAMIGLIDYEAAIHLFKVDKDFEV	480
Qy	421	TAASNIIMSLAVNMLTLLFLTPLFPHYTPPLAVLSAIIIVSAMIGLIDYEAAIHLFKVDKDFEV	480	481	VCM SAYIGVVGSGVEIGLVIATIVISVLRLVLFIA RPRTFVLGNIPNSVIYRNVEHYQNAK	540
Db	421	TAASNIIMSLAVNMLTLLFLTPLFPHYTPPLAVLSAIIIVSAMIGLIDYEAAIHLFKVDKDFEV	480	481	VCM SAYIGVVGSGVEIGLVIATIVISVLRLVLFIA RPRTFVLGNIPNSVIYRNVEHYQNAK	540
Qy	541	HVPGMLILEIDAPIYFANASYLRERITWIDEEBERIKATGETSLQVVIIDMSAVGNIDT	600	541	HVPGMLILEIDAPIYFANASYLRERITWIDEEBERIKATGETSLQVVIIDMSAVGNIDT	600
Db	541	HVPGMLILEIDAPIYFANASYLRERITWIDEEBERIKATGETSLQVVIIDMSAVGNIDT	600	601	SGISMLEEVKKITERRELQVLVNPVSEVMKKLNKSFQNHGKKWYLTVEEAVGACNP	660
Qy	601	SGISMLEEVKKITERRELQVLVNPVSEVMKKLNKSFQNHGKKWYLTVEEAVGACNP	660	601	SGISMLEEVKKITERRELQVLVNPVSEVMKKLNKSFQNHGKKWYLTVEEAVGACNP	660
Qy	661	NLRASKTNPKDDETEGWNV 680		661	NLRASKTNPKDDETEGWNV 680	
Db	661	NLRASKTNPKDDETEGWNV 680				
RESULT 2						
US-09-720-317A-24						
; Sequence 24, Application US/09720317A						
; Patent No. 6696292						
; GENERAL INFORMATION:						
; APPLICANT: Stephen M. Allen						
; APPLICANT: Saverio C. Falco						
; APPLICANT: Catherine J. Thorpe						
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins						
; FILE REFERENCE: BB-1167						
; CURRENT APPLICATION NUMBER: US/09/720.317A						
; CURRENT FILING DATE: 2000-12-21						
; PRIOR APPLICATION NUMBER: 60/092,833						
; PRIOR FILING DATE: 14-07-1998						
; NUMBER OF SEQ ID NOS: 31						
; SOFTWARE: Microsoft Office 97						
; SEQ ID NO 24						
; LENGTH: 658						
; TYPE: PRT						
; ORGANISM: Arabidopsis thaliana						
US-09-720-317A-24						
Query Match 76.0%; Score 2642.5; DB 4; Length 658;						
Best Local Similarity 76.0%; Pred. No. 8.7e-258;						
Matches 501; Conservative 80; Mismatches 75; Indels 3; Gaps 3						
Qy	24	MGSDVYEPPLGMNNFVRHQ-VEVPPQPPFKSLKYSIKETFFPDPLQFKNKPKSKFP	82			
Db	1	MGTEYTFPQGAELHRHHTVEAPQPPFFLKSQYSVKETLFPDPPFQFKNQNASRKP	60			
Qy	83	MLGLQFFFFIFEWAPKTYFQFLKADLIAGITIASLAIPQGISYAKLANLPPILGLYSSTF	142			
Db	61	VVLGXVFLPIFEWAPRYNLKFFKSDLIAGITIASLAIPQGISYAKLANLPPILGLYSSTF	120			
Qy	143	PELLIYAMGSSRDLAGVTAVAGSLMLGSMLSNAVDPNEDPKLYLHLAFTATLPAFVQAA	202			
Db	121	PELLVAVLGRSSRDLAGVTAVASLLTGAMLSKEVDAEDPKLYLHLAFTATTFPAGVLEAS	180			
Qy	203	LGLFRLGLVDFLSHATLIGFMGGNATVVCLOQLKSLGLEHFTHGADIIISVRSVFTOT	262			
Db	181	LGLFRLGLVDFLSHATLIGFMGGNATVVSQQLKIGLFGKHFDTSDTIVISVRSVFSQT	240			
Qy	263	HEWRWESAVLGCVFIFLLSTRYSFKKRPFWFWSAMAPLTSVTLGSLLYFTFHAETHGV	322			
Db	241	HEWRWESVGLGCGFLPFLLSSTYFSIKPKFPFWANAPLTSVTLGSLLYFTFHAETHGV	300			
Qy	323	EVIGELKGLNPPSTNLVFFVSPWMTTAVKGTIVVGIIISLAEGIAVGRSFAMYYKQNNIDG	382			

Db	301	QVIGDLKKGLNPLSGSDLIFTSPYMTAVKTLGITIIAALAEIGAVGRSFAMFKNYNIDG	360
Qy	383	NKEMIAIGTMNVGSGFTSCYLTITGPFPSRAVYNAGCKTAASNIIMSLAVMLTLLFLTPL	442
Db	361	NKEMIAFGMNIIVGSGFTSCYLTITGPFPSRAVYNAGCKTANSNIVMAVMTLLFLTPL	420
Qy	443	PHYTPLVVLSAIIIVSAMLGLIDYBAAIHLFKVDKPDFVVCMSAYIGVVFGSVEIGLVTAI	502
Db	421	PHYTPLVVLSAIIISAMLGLIDYQAAIHLMKVDKPDFLVCNCSAYGVGVFGSVEIGLVAV	480
Qy	503	VISVLRLVLLFTARPTTVLGNIPNSVIYRNVHEYCNQAKHVPQMLTILEIDAPIYFANASYL	562
Db	481	AISLARLLLVSRPKTAVKGNIPNSMIYRNTQEQPSSRTVPGIILEIDAPIYFANASYL	540
Qy	563	RERITRWIDEEERIKATGETSLQVYIIDMSAVGNIDTSGISMLEEVEVKKITERRELOLVL	622
Db	541	RERIIRWIDEEERVKQSGESSLQYIILDM SAVGNIDTSGISMVVEIKKVIDRRALKVLV	600
Qy	623	VNPVSEVMKLNKSKP-QNHLGKWKIYLTVEEAVGACNFNURAKTNPDKDTEGWNV	680
Db	601	SNPKGEVVKLTRSKFIDGHLGKEWMFLTVGEAVEACSYMLHTKTEP-ASKNPEPNV	658
RESULT 3			
US-09-720-317A-26			
; Sequence 26, Application US/09720317A			
; Patent No. 6696292			
; GENERAL INFORMATION:			
; APPLICANT: Stephen M. Allen			
; APPLICANT: Saverio C. Falco			
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins			
; FILE REFERENCE: BB-1167			
; CURRENT APPLICATION NUMBER: US/09/720,317A			
; CURRENT FILING DATE: 2000-12-21			
; PRIOR APPLICATION NUMBER: 60/092,833			
; PRIOR FILING DATE: 14-07-1998			
; NUMBER OF SEQ ID NOS: 31			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 26			
; LENGTH: 646			
; TYPE: PRT			
; ORGANISM: Arabidopsis thaliana			
US-09-720-317A-26			
Query Match 74.1%; Score 2575.5; DB 4; Length 646;			
Best Local Similarity 74.4%; Pred. No. 5e-251;			
Matches 490; Conservative 80; Mismatches 74; Indels 15; Gaps 4			
Qy	24	MGSDVDEYPLGMNFRVHQ-VEVPPQPPPKSKYSLKETFFPDPLRQPKPKASKKF	82
Db	1	MGTEDTYFPQGAELHRHHTVEAPQPQFLKSLQYSVKETLFPDPPRQPKQNASKKF	60
Qy	83	MLGQFPFPFIFEMAPKTYTFQKADLLAGITIASLAIPQGISYAKLANLPILGLYSSFI	142
Db	61	VLGUKYFLPFIEMAPRNLKFFKSDLLAGITIASLAIPQGISYAKLANLPILGLYSSVF	120
Qy	143	PPLIYAMWGSRRDLAVGTAVAGSLLGMSLMSNAVDPNEDPKLYLHLAFTATLFAGVFOAA	202
Db	121	PPLVYAVLGSSRDLAVGTAVASLLTGAMLSKEYDAEKDPKLYLHLAFTATLFFAGVLEAS	180
Qy	203	LGLFRLGLIVDFLSHATIIIGMGGAATVVCLOQLKSLIGLEHFTHGADIIISVMRSVFQT	262
Db	181	LGI FRLGFI VDFLSHATIVGPMGAATVWSLQQKGFGLKHFDTSDTIVSVMRSVFSQT	240
Qy	263	HEWRWESAVLGCVFIFFLLSTRYFSKCRPRFPWVSAMAPLTSVILGSLLVYFTHAEKHGV	322
Db	241	HEWRWESVLCGCGFLFPFLLSTRYFSIKKPEFPWVAMAPLTSVILGSLLVYFTHAEKHGV	300
Qy	323	EVIGELKKGLNPPSLTNLVFVSPYMTAVKTVGVVIGIISLAEGIAVGRSFAMFKNYNIDG	382
Db	301	QV-----GSDLIFTSPYMTAVKTLGITIIAALAEIGAVGRSFAMFKNYNIDG	348

2



QY 383 NKEMIAIGTMNVGSFTSCYLTTPGPFSSAVNNAGCKTAASNIIMSLAVMLTLLFLTPL 442  
Db 349 NKEMIAFGMNNI VGSFTSCYLTTPGPFSSAVNNAGCKTAMSNIVMAIAMFTLLFLTPL 408  
QY 443 FHTPLVLSAIIISAMGLGIDYEAATHLFKVDKDFVVCMSAYIGVVGFSVEIGLVIAI 502  
Db 409 FHTPLVLSAIIISAMGLGIDYQAALHLKVDKDFVLCMSAYGVVGFSVEIGLVVAV 468  
QY 503 VISLRLVLLFIARPTFVLGNIPNSVIYRNVEHYQNAKHVPGMLIILSIDAPIYFANASYL 562  
Db 469 AISIARLULLFVRPKTAVKGNIPNSMIYRTEQPPSSRTVPGIILSIDAPIYFANASYL 528  
QY 563 RERITRWIDEEERIKATGETSLOYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVL 622  
Db 529 RERIIRWIDEEERVKOGSSSLQYIILDMASVGNIDTSGISMMVEIKKVIDRRLKVL 588  
QY 623 VNPVSEVMKKLNKSKF-QNHGLKKKIYLTVEEAVGACNFNLRAKTNPKKDETEGNNV 680  
Db 589 SNPKGEVVKKLTRSKFIDGHLGKEMELTVGEAVEACSYMLHTFKTEP-ASKNEPNNV 646

RESULT 4  
US-09-720-317A-28  
; Sequence 28, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 28  
; LENGTH: 646  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-720-317A-28

Query Match 66.6%; Score 2314; DB 4; Length 646;  
Best Local Similarity 67.2%; Pred. No. 1.3e-224;  
Matches 430; Conservative 104; Mismatches 122; Indels 4; Gaps 2;

QY 42 HOVEVPPQPFKSLKYSKATETFPDDPLQFKNK-PASKKFMGLQFPFPIFEWAPKYT 100  
Db 10 HOVEIPPPQFLSKNLTNEILFADDPFRIRNESKTSKKIELGRHVFPFILEWARGYS 69  
QY 101 FQFLKADLIAGITIASIAIPQGISYAKLANLPILGLYSSFIPLIYAMGSSRDLAAGT 160  
Db 70 LEVLSKDVISGIIIASIAIPQGISYAKLANLPILGLYSSVLPVLYAINGSSRDLAGT 129  
QY 161 VAGSLMGLMNLNAVDNEDPKLYLHLATATLFAGVFOAALGLFRLGLIIVDFLSHATI 220  
Db 130 VAVASLLTAAMLGKEVNAVNPVKLYLHLATATLFAGLMOTCLGLRLGFWVBILSHAAI 189  
QY 221 IGFWGGAAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQTHEWRESAVLGCVFIFL 280  
Db 190 VGFWGGAAATVVCLOQLKGLLGLHHFTHTDITVLRISIFSQSHRWRESVGLGCCFLIFL 249  
QY 281 LSTRYFKKPRFVFWSAMAPLTSVILGSLVAVFTTHAEKKGVEVIGELKKGLNPPSLTNL 340  
Db 250 LTTKYISKPKPLZWISAMSPVSVIQTIFLFLHDQFHGIQFIELKKGINPPSIHL 309  
QY 341 VFVSPYMTTAVKGIIVGIIISLAEGIAVGRSFAMKYKNYIDGNKEMIAICTMNVGSGFTS 400  
Db 310 VFTPPYMLALKVGIITGVIALAEGIAVGRSFAMKYKNYIDGNKEMIAFGMMNIGSFSS 369  
QY 401 CYLTTGPFSSAVNNAGCKTAASNIIMSLAVMLTLLFLTPLPHYPLVLSAIIIVSAML 460

Db 370 CYLTTGPFSSAVNNAGCKTALSNNVMAVAVAVTLLFLTPLFFYTPLVVLSIIIAAML 429  
QY 461 GLIDYEAATHLFKVDKDFVVCMSAYIGVVGFSVEIGLVIAIIVISLRLVLLFIARPTFV 520  
Db 430 GLVDYEAATHLWKLDKDFDFVCLSAVLGVVFGTIEIGLILSVGISVMRLVLFVGRPKIYV 489  
QY 521 LGNIPNSVIYRNVEHYQNAKHVPGMLIILSIDAPIYFANASYLBERITRWIDEEERIKAT 580  
Db 490 MGNIQNSEYRIEYHYPQAITRSSLLIHLIDGFIYFANSTYLRDRIGRWIDEEEDKLRTS 549  
QY 581 GETSLOYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNKSKFQ 640  
Db 550 GDISLOYIIVLDMASVGNIDTSGISMLEEELNKILGRRELKLVIANPGAEVMMKLSKSTFIE 609  
QY 641 HLGKKKIYLTVEEAVGACNFNLRAKTNPKKDETEGNNV 680  
Db 610 SIGKERIYLTVAEAAACDFMLHTAKPDSVPVE---FNNV 646

RESULT 5  
US-09-720-317A-4  
; Sequence 4, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 4  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-720-317A-4

Query Match 62.3%; Score 2164; DB 4; Length 579;  
Best Local Similarity 71.0%; Pred. No. 1.6e-209;  
Matches 403; Conservative 86; Mismatches 79; Indels 0; Gaps 0;

QY 105 KADLIAGITIASIAIPQGISYAKLANLPILGLYSSFIPLIYAMGSSRDLAAGTAVAG 164  
Db 3 ESDLIAGITIASIAIPQGISYAKLANLPVGLYSSFPVPLVYALMGSSKDLAAGTAVAG 62  
QY 165 SLIMGSLMNLNAVDNEDPKLYLHLATATLFAGVFOAALGLFRLGLIIVDFLSHATIIGFM 224  
Db 63 SLLISSMLGSEVSPTEPNVLYLHLATATLFAGVFOASLGLRLGLFIVDLLSHATIIGFM 122  
QY 225 GGAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQTHEWRESAVLGCVFIFELLSTR 284  
Db 123 AGAATVVCLOQLKGLMGLVHFTTSDTVSVMSVSPSQTHQWRWESVLLCGGFFLLVTR 182  
QY 285 YFSKKKPRFVFWSAMAPLTSVILGSLVAVFTTHAEKKGVEVIGELKKGLNPPSLTNLVFVS 344  
Db 183 FISKRPKLFWISAAAPLTSVILGSLVAVLTHAENHIGIEVIGYLLKGLNPPSVTSLOFSP 242  
QY 345 PYMTTAVKGIIVGIIISLAEGIAVGRSFAMKYKNYIDGNKEMIAICTMNVGSGFTSCYLT 404  
Db 243 PYMMLALKGTIITGVIALAEGIAVGRSFAMKYKNYIDGNKEMIAICTMNVLSGLTSCYLT 302  
QY 405 TGPFSRSVNNAGCKTAASNIIMSLAVMLTLLFLTPLPHYPLVLSAIIIVSAMGLID 464  
Db 303 TGPFSRSVNNAGCKTAMSNVMSLAVMVMTLLFLTPLPHYPLVLSAIIIVSAMGLVD 362  
QY 465 YEAAIHLKVDKDFVVCMSAYIGVVGFSVEIGLVIAIIVISLRLVLLFIARPTFVLGNI 524  
Db 363 FGNALHLWRVDKDFVCAGAYLGVVFGSVGLVAVAVSLRLVLLFVARPTTVLIGNI 422



QY 581 GETSLQYVIIDMSAVGNIDTSGISMLBEVKKITERRELOQLVLPVNPVSEVMKKLNKSKFON 640  
Db 543 KHSQQLIIEEMSAVSGVDNGVSFFKELKKTAKKDIELVFNPLSEVVEKLRQADEQK 602  
QY 641 H-LGKKWIYLTVEEAVGA 657  
Db 603 EFMPEPFLFJTVAEAVAS 620

RESULT 8  
US-09-720-317A-2  
; Sequence 2, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 688  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-720-317A-2

Query Match 53.1%; Score 1843.5; DB 4; Length 688;  
Best Local Similarity 56.2%; Pred. No. 5.2e-177;  
Matches 354; Conservative 115; Mismatches 156; Indels 5; Gaps 3;

QY 41 VHOVEVPPQPFKSLKYSKLETFFDDPLRQFNKPASKKFMGLQGFPPPIFEWAPKYT 100  
Db 63 VHKVAPPARSTASKMKVRKETFFDDPPFAFKGPQGTQWLMAVRYLEPILDWPSYS 122  
QY 101 FQFLKADLIAGITIASLAIPOGISYAKLANLPPILGLYSFFIPPLIYAMGSSRDIAVGT 160  
Db 123 LSLFKSDLVAGLTIASLAIPOGISYAKLASLPIILGLYSFVPVYAVLGSSRDIAVGP 182  
QY 161 VAVGSLLMGSKLNAVDPNEDPKLYLHATATLPAFQVFAALGLFRLGLIVDFLSHAT 220  
Db 183 VSISLLMGSLRAQVSPTEPTLFLQLAFTSTLPAFLVQASGLILRLGFVIDFLSKATL 242  
QY 221 IGFMGGAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQTHEWRWESAVLGCVFIFL 280  
Db 243 VGFMAAGAAITIALQQLKYVLGIANFTRKTDIVSVMESVWRSVHHGWNKQTIIVGVSLVF 302  
QY 281 LSTRYFSKKRPRFPFWSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGLNPPSLTNL 340  
Db 303 LSARHVSIRPKLFWSACAPLASVITSLVLFLFKAQNHGIIIGOLKGLNRPSSDKL 362  
QY 341 VFVSPYMTTAVKTVGVVGIISLAEGIAVGRSPFAMYKNYNIDGNKEMIAIGTMNVVSGFTS 400  
Db 363 LFDATVILGLTKWTKGLTVGTIISLLEGIAVGRFTASLKDQYDGNKEMWIAIGTMNVVSGCTS 422  
QY 401 CYLTTGPFPSRANVYAGCKTAASNIIMSLAVMLTLLFLTPLFHYTPLVLSAIIVSAML 460  
Db 423 CYVTGTAFPSRANVHNAAGCKTASNVIMALTVMVTLFLFPLFVYTNVNVILGAIITAAVI 482  
QY 461 GLIDYEAAILHFKVDKDFVVCMSAYIGVVGVSVEIGLVIAIVISVLRLVLLFIARPTFV 520  
Db 483 GLIDFPVAVHIWKWDKMDFLVCVAPAGVIFISVQEGLAIVAGISIFRVLQMQITRPMV 542  
QY 521 LGNIPNSVIYRNVEHYQNAKHVPGMLILETDPIYFANASYLRLERTRWIDEEERIKA 580  
Db 543 QGNIKGTDIYRDLHHYKEAQRVSGFLILATEAPINFANSYLNERNIKRWI--BEESFEQD 600  
QY 581 GETSLQYVIIDMSAVGNIDTSGISMLBEVKKITERRELOQLVLPVNPVSEVMKKLNK-SKFO 639

Db 601 KHTELHFIILDSAVPAIDTSGIAFLIDIKSIEKGLLELVLPVPTGEVMEKIQRANEAE 660  
QY 640 NHLGKKWIYLTVEEAVGACNFNLRSKTNP 669  
Db 661 NYFRPDCLYLTGGEAI--ASLSALAKWTKP 688

RESULT 9  
US-09-720-317A-20  
; Sequence 20, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 20  
; LENGTH: 656  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-09-720-317A-20

Query Match 51.8%; Score 1798.5; DB 4; Length 656;  
Best Local Similarity 53.6%; Pred. No. 1.7e-172;  
Matches 332; Conservative 123; Mismatches 163; Indels 1; Gaps 1;

QY 41 VHOVEVPPQPFKSLKYSKLETFFDDPLRQFNKPASKKFMGLQGFPPPIFEWAPKYT 100  
Db 30 VYKVGYPKPLATEFTETLRSTFFHDNPLRQYKQSGPRFRFMGLEFLFPFGWGRDYS 89  
QY 101 FQFLKADLIAGITIASLAIPOGISYAKLANLPPILGLYSFFIPPLIYAMGSSRDIAVGT 160  
Db 90 LNKFKGDLIAGLTIASLAIPOGISYAKLANLPPQGLYSFFIPPLIYAMGSSRDIAIGP 149  
QY 161 VAVGSLLMGSKLSNAVDPNEDPKLYLHATATLPAFQVFAALGLFRLGLIVDFLSHAT 220  
Db 150 VAVSLLIGSLLOAEVDHVNKEENRMTATFFAGITQAAALGLFRLGLFIEFLSHAAL 209  
QY 221 IGFMGGAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQTHE-WRWESAVLGCVFIF 279  
Db 210 VGFMGGAATIALQQLKYVLGIANFTRKTDIVSVMESVWRSVHHGWNKQTIIVGVSLVF 269  
QY 280 LSTRYFSKKRPRFPFWSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGLNPPSLTN 339  
Db 270 LLFAKYIGKKRKLFWVPAIAPISVILATFFVYITRADKQGVQIVKHIEQGINPSSVHK 329  
QY 340 LVFVSPYMTTAVKTVGVVGIISLAEGIAVGRSPFAMYKNYNIDGNKEMIAIGTMNVVSGFT 399  
Db 330 IYFTGFVAKGPKIGVVGIGVGLTEAVAIGRTFAAMKQYQLDGNKEMVALGTNIVGSM 389  
QY 400 SCYLTGPFPSRANVYAGCKTAASNIIMSLAVMLTLLFLTPLFHYTPLVLSAIIVSAM 459  
Db 390 SCYVTTGFSRANVFMACCKTPVSNVMSVVLLTLLVITPLFKYTPNAILGSIISAV 449  
QY 460 LGIDYEAAILHFKVDKDFVVCMSAYIGVVGVSVEIGLVIAIVISVLRLVLLFIARPTF 519  
Db 450 IGLVDYEAAILWKVDKDLFIACMGAPFGVVFVEIGLLIAVAISFAKILLQVTPRTA 509  
QY 520 VLGNIPNSVIYRNVEHYQNAKHVPGMLILETDPIYFANASYLRLERTRWIDEEERIKA 579  
Db 510 LUGNPFTIYRNISQYPEAKLTGPVIVRVDSYFNSNVRERILLRWLTDDEERAKA 569  
QY 580 TGSETSQQYVIIDMSAVGNIDTSGISMLBEVKKITERRELOQLVLPVNPVSEVMKKLNKSKFO 639  
Db 570 VGLPKISFLIVEMSPVIDIDTSGIHALEDLYKNLQKQMDQLILSNPGSVVIEKQLQASKLT 629

```
Qy 640 NHLGKKWIYLTVEEAVGAC 658
Db 630 EHISSNIFLAVSDAVRC 648

RESULT 10
US-09-720-317A-30
; Sequence 30, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Hordeum vulgare
US-09-720-317A-30

Query Match 50.7%; Score 1760.5; DB 4; Length 660;
Best Local Similarity 53.0%; Pred. No. 1.2e-168;
Matches 326; Conservative 127; Mismatches 161; Indels 1; Gaps 1;

Qy 42 HOVEVPPQPFKSLKYLKETFPPDDPLRQFNKPKASKKFMGLQLOFFPPIPEWAPKYTF 101
Db 35 YKVGFPKAGVFAEFAEGVKETFDADPLEYDQDPRSKLWLSVHLFPVLDMSYTF 94
Qy 102 QFLKADLIAGITIASLAIPQGISYAKLANLPPIILGYSSFPPLIYAMGSSRDLAGVT 161
Db 95 GKFKGDLVAGLTIASLIPQDQIGYAKLANLPQHLGYSSFPPLIYALMGSSRDIAIGPV 154
Qy 162 AVGSLMGSLNAPVDPNPKLYLHAFATLTFAGVFQAAALGLFRLGLIVDFLSHATII 221
Db 155 AVSLLLGLTLLQBEIDPVKNPLSYSLAFTATTFAGITQAMLGFFRLGFTIEFLSAAIV 214
Qy 222 GFMGGAATVVCLOQLKSLGLEHFTHGADIISVMRSVFTOT-HIEWRWSAVLGCVFIFL 280
Db 215 GFWAGAAITIALQOLKGLLGIKFTKSDIISVMESVGNVQHWKQOTILIGSSFLAFL 274
Qy 281 LSTRYFSKKRPRFFWVSAMAPLTSVLGSLLYFTTHAEKHGVEVIGELKGLNPPSLTNL 340
Db 275 LTTKYIAKNNKLFWSAIAPLISVISTPCVYITRADKQGVAVKNIKQGINPPSFDLI 334
Qy 341 VFVSPYMTTAVKTGVVGIISLAEGTAVGRSFAMYNKYNIDGNKEMIAIGTMNVVGSPTS 400
Db 335 YWGGPIYAKGFRIGVWGSWVALTEAIGTKTFAAMKDQYIDGNKEMVAGLTGMNIVGSMTS 394
Qy 401 CYLTGPFBSRSAYNNAGCKTAASNIIMSLAVMLTLLFTPLPHYTPLVVLSAIIVSAML 460
Db 395 CYVATGSFSRSAYNNAGCKTASNVVMAIVMLTLLTLLTPFKYTPNALIASIINAVV 454
Qy 461 GLIDYEAAILFKVDKDFVVCMSAIVGVVFGSVEIGLVAIVSVLRLFTARPTFV 520
Db 455 NLVDYETAYILWVKDMDFVALLGAFPGVGVFASVEYGLIATAISLGLKILLQVTRPTAL 514
Qy 521 LGNI PNSVIYRNVEHYQNAKHVPGLMLLEIDAPIYFANASYLRERITRWIDEEERI KAT 580
Db 515 LGNLPTTIRYRVEQPEATKVPGLMVRVDSAIYFTNSNYKERTLRWLURDEEEQOEQ 574
Qy 581 GETSLQYVIDMSAVGNIDTSGISMLEEVKKITRERELQVLVNPVSEVMKKLNKSKFQ 640
Db 575 KLSKTEFLVELSPVTDIDTSGIHALEELKALEKRIQLILANPGPAVIQKLSAKFTD 634
Qy 641 HLGKKWIYLTVEEAV 655
Db 640 NHLGKKWIYLTVEEAV 655
```

```
Db 635 LIQDDKIFLSVGDAV 649

RESULT 11
US-09-720-317A-23
; Sequence 23, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Sporobolus stapfianus
US-09-720-317A-23

Query Match 49.6%; Score 1723.5; DB 4; Length 660;
Best Local Similarity 55.8%; Pred. No. 6.5e-165;
Matches 349; Conservative 99; Mismatches 160; Indels 17; Gaps 5;

Qy 41 HOVEVPPQPFKSLKYLKETFPPDDPLRQFNKPKASKKFMGLQLOFFPPIPEWAPKYT 100
Db 35 VHKWVPPPTASTASKLTKRLKETFPPDDPFRGFGQOPARVQWVLAVKYLPFLDMLPAYS 94
Qy 101 FQFLKADLIAGITIASLAIPQGISYAKLANLPPIILGYSSFPPLIYAMGSSRDLAGVT 160
Db 95 LSLFSDSLIAGITIASLAIPQGISYAKLANLPPIILGYSSFPPLIYAVLGSSRDLAGVP 154
Qy 161 VAVGSLMGSLNSNA---VDPNEDPKLYLHAFATLTFAGVFQAAALGLFRLGLIVDFLSH 217
Db 155 VSISSLINGPCASQPHCGADAVPAARLH---ATLPAGIFQASLGIURLGLFIIDFLSK 210
Qy 218 ATIIIFMGGAATVVCLOQLKSLGLEHFTHGADIISVMRSVFTOTHEWRWSAVLGCVF 277
Db 211 ATLGVFMAGAAIIVSLQQLKALLGIVHFTTEMGIVPVNASVFHHTKEMSWQITLMGVCEL 270
Qy 278 FLLSTRYFSKKRPRFFWVSAMAPLTSVLGSLLYFTTHAEKHGVEVIGELKGLNPPSL 337
Db 271 VFLLVARHVSIRWPRLFWVSACAPLVSIVISTLVVFLFKAQNHGISIIQOLKGLNRP 330
Qy 338 TNLVFSVPMYTTAVKTGVVGIISLAEGTAVGRSFAMYNKYNIDGNKEMIAIGTMNVVGS 397
Db 331 DKTNIDTTLGLTMTKLTGLVTGIIISLTEGIAVGRTFASLKEYOIDGNKEMVAGLTGMN 390
Qy 398 FTSCLYLTGPFBSRSAYNNAGCKTAASNIIMSLAVMLTLLFTPLPHYTPLVVLSAIIVS 457
Db 391 CTSCTVTTGAFSRSVPVNHAGCKTASNVIMALTVMVTLFLMPLFVYTPNVVLGAIITA 450
Qy 458 AMGLIDYEAAILFKVDKDFVVCMSAIVGVVFGSVEIGLVAIVSVLRLFTARPR 517
Db 451 AVIGLIDIPAVTHIWKMDKMDFLVCVCAPAGVLFISVQEGLAIAVGISVFRVLLQITRPK 510
Qy 518 TFLVGNIPNSVIYRNVEHYQNAKHVPGLMLLEIDAPIYFANASYLRERITRWIDEEERI 577
Db 511 ITVQGNIMGTDIYRNHLHQYKDAQIRPGFLILATEAPINFANSYLNRIKRWI--EEESS 568
Qy 578 KATGETSLQYVIDMSAVGNIDTSGISMLEEVKKITRERELQVLVNPVSEVMKKLNKSK 637
Db 569 AQTQKTELFRVLDLSAVPAIDTSGVAFLLIDIKKSIKRGLEGLVLPNPTGEGHKNNTASE 628
Qy 638 -----FQNHGKKWIYLTVEEAVGA 657
Db 629 RGTQAPQVGIA---CILITGEAVS 650
```

## RESULT 12

US-09-720-317A-25  
; Sequence 25, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 25  
; LENGTH: 644  
; TYPE: PRT  
; ORGANISM: Stylosanthes hamata  
US-09-720-317A-25

Query Match 43.3%; Score 1503; DB 4; Length 644;  
Best Local Similarity 47.0%; Pred. No. 1.2e-142; Mismatches 124; Indels 22; Gaps 6;  
Matches 301; Conservative 124; Mismatches 193; Indels 22; Gaps 6;  
Qy 33 LGMNFRVHQ--VEVPPPPQPFKSLKYSKETFPPDDPLRQFNKPKASKKFMGLQF-- 88  
Db 4 LGTEQFSERQWLNPNPLTKKFLGPLX-----NKFTSSSSKKEKTRVSVFLA 55  
Qy 89 -FPPIFWAPKYTOFLKADLIAGITIASIAPOGISYAKLANLPPIILGYSGFIPPLIY 147  
Db 56 SLFPILSWIRYSATKPKDGLLSGLTLASLSIPQSIGYANLAKLDPQYGLYTSVIPVIV 115  
Qy 148 AMGSSDPLAVGTAVGSSLMGSMNAVDNEDPKLYLHLATATLPAQVFOAALGLFR 207  
Db 116 ALMGSSREIAGPVAVVMSLLSVKVIDPDAPHNDYRLNLFVTVLFAIGIFOTAFGLR 175  
Qy 208 LGLIVDFLSHATITGFMGGATVVCLOQLKSLGLEHFTHGADIISVMRSVFTQTHE-- 264  
Db 176 LGFLVDFLSHAALVGFMAAGIIVGLQKLLGLTHFTTKTDAVAVKSVYTSLHQOIT 235  
Qy 265 ----WRWESAVLGCVFIFLLSTRYSKGRPRFFWVSAMAPLTSVILGSLLVYFTHAEKH 320  
Db 236 SSENWSPNLNFVIGCSFLIFLAARFICGRNKKFFWLPALAPLLSVILSTLIVFLSKGDKH 295  
Qy 321 GVEVIGELKGLNPPSLTNLVFVSPYMTTAVKTIIVVGIISLAEGIAVGRSFAMYKYN 380  
Db 296 GWNIIKHVQGLNPSVHKLQNLGPHVGOAKIGLSAIIATLTAIAVGRSFANIKGYHL 355  
Qy 381 DGNKEMIAIGTMNVVSGFTSCYLTGTPFSRANVYAGCKTAASNIIMSLAVMLTLFL 440  
Db 356 DGNKEMLAMGNMIAAGSLTSCYVSTGSRFATNVSAGCKTANVIMVAVTLCLLEFT 415  
Qy 441 PLFHYTPLVLSAIVSAMLGLIDYEAAILFKVDKDFVVCMSAYIGVVFGSVGLVI 500  
Db 416 RLLYTPMIALSIIISALPGLIDIGEAYHWKVDKDFLACLAGAPFGLVFSIEIGLLI 475  
Qy 501 AVISVLRLVLLFIARPTFLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDA-PIYANA 559  
Db 476 ALSISFAKILLQAIIRPGVGLGRIPPTTEAYCDVAQYPMVTTFGILVIRISSGLCPANA 535  
Qy 560 SYLRERITRWI--DEEBERIKATGTSLOYYIDMSAVGNIDTSGISMLEBVKKITERREL 618  
Db 536 GFVRERILKWEDEQONIEAAKGRVQAIIDMTDLTNVDTSGILALELHKLLSRGV 595  
Qy 619 QLVLVNPSVEMKXKLNKSKFQNLHGKWKWILYTVBEAVGAC 658  
Db 596 ELAMVNRPRWEVIHKLKVANFVDKIGKERVFLTVAEAVDAC 635

## RESULT 13

US-09-720-317A-22  
; Sequence 22, Application US/09720317A

US-09-720-317A-31  
; Sequence 31, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 31  
; LENGTH: 685  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-720-317A-31

Query Match 28.7%; Score 997; DB 4; Length 685;  
Best Local Similarity 34.0%; Pred. No. 1.8e-91;  
Matches 217; Conservative 140; Mismatches 232; Indels 50; Gaps 13;  
Qy 47 PPQPFPKSLKYSKETFPPDDPLRQFNKPKASKKFMGLQFPPFIFEWAPKYTF-QFLK 105  
Db 51 PPSIPE-----DDIFSGWTAKIKRMRLVDWIDTLFPCFRWIRTYRWSEYFK 96  
Qy 106 ADLIAGITIASIAPOGISYAKLANLPPIILGYSSIPPIIYAMGSSRDLAVGTAVGS 165  
Db 97 LDLMAGITIGIMLVPOAMSAYAKLAGLPPYIGYSSFPVFPVYAFIAGSSRQLAIGPVALS 156  
Qy 166 LMGSMLSNAVDNEDPKLYLHLATATLPAQVFOAALGLFLGLIVDFLSHATITGFMG 225  
Db 157 LLVSNALGGIADNNE--LHIELAILLALLVGLIEGIMLLRGLWIRFISHSVISGFTS 214  
Qy 226 GAATVVCLOQLKSLGLEHFTHGADIISVMRSVFTQTHEWRWESAVLGCVFIFLLSTRY 285  
Db 215 ASAVITGLSKYKIFLGYIS-IARSSKIVPIVESIAGADKQFQWPPFVMSGLILVILQVMKH 273  
Qy 286 FSKRPRFFWVSAMAPLTSVILGSLLVYFTHAEKHGVEVIGELKGLN----PPLTNLV 341  
Db 274 VGAKKELQFLRAAAPLTGIVLGTAKVPHPP--SISLVGEIPOGLTTFSPRSPDHAK 331  
Qy 342 FVSPYMTTAVKTIIVVGIISLAEGIAVGRSFAMYKYNIDGNKEMIAIGTMNVVSGFTSC 401  
Db 332 TLLP--TSALITG-----VPILSVGIKALAAKNRYELDSNDFLGLGVANILGSLFSA 384  
Qy 402 YLTGTPFSRANVYAGCKTAASNIIMSLAVMLTLFLTPLFHYTPLVLSAIVSAMLG 461  
Db 385 YPATGFSRSRANVNESEAKTGLSLITGIIIGCSLLFLTPMPKYIPOCALAAIVISAVSG 444  
Qy 462 LIDYEAAILFKVDKDFVVCMSAYIGVVFGSVGLVIAIVISVLRLVLLFIARPTFV 520  
Db 445 LDYDRAIFLWRVDRKDFSLTITLITLFFG-IEIGLVGVGFSLAFVHESANPHIAV 503  
Qy 521 LGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYANASYLRERITRW----IDE----- 572  
Db 504 LGRLPQTTVYRNIIKQYPEAYTYNGIVIVRIDSPIYFANISYIKORLEVEVAVDKYNRG 563  
Qy 573 -EERIKATGTSLOYYIDMSAVGNIDTSGISMLEBVKKITERRELQVLVNPVSEVMK 631  
Db 564 LEVDRI-----NFVILEMSPVTHIDSSAVEALKELYQYKTRDITQLAISPNKDVHL 615  
Qy 632 KLNKSKFQNLHGKWKWILYTVBEAVGACNPLRASKTNPK 670  
Db 616 TIARSGHVELVGKBNFFVRVHDVAVQVCLOYVOSSNLEDK 654

```

, Patent NO. 6696292
, GENERAL INFORMATION:
, APPLICANT: Stephen M. Allen
, APPLICANT: Saverio C. Falco
, APPLICANT: Catherine J. Thorpe
, TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
, FILE OF INVENTION: BB-1167
, CURRENT APPLICATION NUMBER: US/09/720,317A
, CURRENT FILING DATE: 2000-12-21
, PRIOR APPLICATION NUMBER: 60/092,833
, PRIOR FILING DATE: 14-07-1998
, NUMBER OF SEQ ID NOS: 31
, SOFTWARE: Microsoft Office 97
, SEQ ID NO 22
, LENGTH: 593
, TYPE: PRT
, ORGANISM: Triticum aestivum
, US-09-720-317A-22

```

Query Match	27.7%	Score	961.5	DB	4	Length	593
Best Local Similarity	35.0%	Pred.	No.5.5e-88				
Matches	204	Conservative	127	Mismatches	235	Indels	17
	Gaps						
Qy	91	PIPEWAKYTF-QFLKADLIAGITIASIAIPQGISYAKLANLPILGLYSSFIPPLVAM	149				
Db	22	PCLAWMSYRKWEDFQADLAAGITVGNLVFQAMSYAKLAGLHPYIGLYTGFVFLVVAI	81				
Qy	150	MGSSRDVAGTVAVVAGSLIMGSMLSNADPNPDPKLYLHLAFTATFLFAGVFOAALGLFRLG	209				
Db	82	FGSSRQLAVGVPVALVSLVSNVLGGIV--NSSSELYTELAITLLAFWVGILCELMALLRLG	139				
Qy	210	LIVDFLSHATIIIPMGGAATVVCLOQLKSILGLHFHTGADIISVMRSVFTOTHEWRWES	269				
Db	140	WLIRFISHSVSGTFTTASIVIGISQKIFYLGS-VTRSSKIIPLIESIAGIDQFGWPP	198				
Qy	270	AVLGCVFIFLLSRYPFSKRRPRFWMVSAMAPLTSVILGSLLVYFTHAEKHGVIVIGELK	329				
Db	199	FVMSGAFVLILLIMKLGKTNKKLRFLRASGPLTAVVLGTLFVKIFRPT--AIVSGEIP	256				
Qy	330	KGLNPPSLTNLVFVSP-----YMTAVKTGIVGIIISLAEGTAVGRSFAMYKNYINDGKE	385				
Db	257	QGL--PS-----FSIPRGFEHLMSPMTAILITGVAILLESVGIKAKALAAKNGYELDSNKE	309				
Qy	386	MIAGTWMNVGSFTSCVLTTPGFRSRVAVNYNAGCKTAASNIIMSLVMTLLFTPLPHY	445				
Db	310	LFGLGLSNI CGSFFSAPATGFSRFSRVAHNHESGAKTGLSGIIMGIIICSAALLFWPIETD	369				
Qy	446	TPLVVLISAIIVSAMGLIDYEAAIHLFKVDKDFWCWMSAYIGVVGVSVEIGLVAIVIS	505				
Db	370	IPQCALAAIVISAVTGLVDYEEAIFLWGIDKKDFLWAMTFTTTTLTFGIEIGLVGVGFS	429				
Qy	506	VLRVLLTAPRTPFVLGNI PMSVIVRNVEHYQNAKHVPGLMLILEIDAPIFANNASYLRER	565				
Db	430	LAFVIHESAMPHIAVGLRPLGTTVYRNTLQYPEAYTYNGIVVVRVDAPVPANTSYIKDR	489				
Qy	566	ITRWIDEEERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITRERELQVLVNP	625				
Db	490	LREYELKLPNSNRGPDVGRVYFVILEMSPVTYIDSSAVQALKDLHOEYKARDIQIATNP	549				
Qy	626	VSEVWKKLNKSKFNHGLKKWIIYLTVEEAVGACNFNLRAKTN	668				
Db	550	NRQVHLLSRAGIIDMTGAGCFCFVRVHDADVOVCLQHVRSRSSN	592				

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2005, 09:39:29 ; Search time 56 Seconds  
(without alignments)  
4771.247 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLRNTI.....NLRASTNPKDTEGNNV 680

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpa/US06\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpa/US11\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/1/pubpa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2330.5	67.1	657	US-10-437-963-140667	Sequence 140667
2	2015	58.0	414	US-10-424-599-273958	Sequence 273958
3	1853.5	53.3	695	US-10-425-115-235264	Sequence 235264
4	1850	53.2	659	US-10-424-599-207945	Sequence 207945
5	1811.5	52.1	666	US-10-437-963-198336	Sequence 198336
6	1773.5	51.0	662	US-10-437-963-179488	Sequence 179488
7	1762	50.7	681	US-10-425-115-260426	Sequence 260426
8	1733.5	49.9	656	US-10-437-963-175893	Sequence 175893
9	1724	49.6	397	US-10-424-599-273959	Sequence 273959
10	1723.5	49.6	653	US-10-437-963-179495	Sequence 179495
11	1722.5	49.6	662	US-10-425-115-366870	Sequence 366870

12	1563.5	45.0	659	US-10-437-963-174545	Sequence 174545
13	1500	43.2	653	US-10-424-599-195017	Sequence 195017
14	1460	42.0	575	US-10-425-114-69272	Sequence 69272, A
15	1436	41.3	405	US-10-425-115-228851	Sequence 228851
16	1406	40.5	640	US-10-425-115-224859	Sequence 224859
17	1379.5	39.7	687	US-10-424-599-217235	Sequence 217235
18	1344.5	38.7	587	US-10-437-963-180840	Sequence 180840
19	1222	35.2	490	US-10-424-599-225979	Sequence 225979
20	1197.5	34.5	730	US-10-437-963-178936	Sequence 178936
21	1185.5	34.1	543	US-10-424-599-217231	Sequence 217231
22	1018.5	29.3	330	US-10-424-599-282576	Sequence 282576
23	1013	29.2	392	US-10-424-599-197192	Sequence 197192
24	1011.5	29.1	339	US-10-424-599-273793	Sequence 273793
25	998	28.7	485	US-10-739-930-9381	Sequence 9381, Ap
26	994	28.6	485	US-10-424-599-203123	Sequence 203123
27	975.5	28.1	392	US-10-425-115-260428	Sequence 260428
28	955.5	27.5	689	US-10-437-963-195183	Sequence 195183
29	923	26.6	271	US-10-767-701-43225	Sequence 43225, A
30	921	26.5	224	US-10-767-701-32265	Sequence 32265, A
31	905	26.0	323	US-10-767-701-40999	Sequence 40999, A
32	887	25.5	303	US-10-437-963-188282	Sequence 188282
33	833.5	24.0	522	US-10-437-963-109890	Sequence 109890
34	794	22.8	156	US-10-424-599-175150	Sequence 175150
35	784.5	22.6	187	US-10-424-599-175149	Sequence 175149
36	737.5	21.2	279	US-10-424-599-161259	Sequence 161259
37	730.5	21.0	280	US-10-424-599-278772	Sequence 278772
38	715	20.6	262	US-10-425-115-260430	Sequence 260430
39	697	20.1	265	US-10-424-599-246655	Sequence 246655
40	674	19.4	227	US-10-425-115-366867	Sequence 366867
41	668	19.2	714	US-09-749-589-4	Sequence 4, Appli
42	668	19.2	714	US-10-684-532-4	Sequence 4, Appli
43	662	19.1	751	US-09-795-693-14	Sequence 14, Appl
44	662	19.1	751	US-10-156-239-14	Sequence 14, Appl
45	662	19.1	751	US-10-199-485-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1  
US-10-437-963-140667  
; Sequence 140667, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 140667  
; LENGTH: 657  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_41845C.1.pep  
US-10-437-963-140667

Query Match 67.1%; Score 2330.5; DB 16; Length 657;  
Best Local Similarity 70.6%; Pred. No. 8.4e-193;  
Matches 441; Conservative 83; Mismatches 100; Indels 1; Gaps 1;  
QY 43 QVEVPPQFFKSLKYSLETFFDDPDLRQF-KNKPASKFKMLGLQFFFPFIPWPKYTF 101  
DB 18 RVMPDAKPFLETLGGNMKETFLPDDFFRVRRRCGRRAAALRYVFFPFWAHSYTL 77





```
Db 370 YFTGPIATGFKGIVAGMIGLTERAIGKRTFAALKDYRIDGKEMMALGTMMIVGSLTS 429
Qy 401 CYLTGTPFSRANYNAGCKTAASNIIMSLAVMLTLLFLPLFHYTPLVVLISAILVSAML 460
Db 430 CYVATGSFSRANYNAGCKTAVSNVVMVMTLLLTITPLFKYTNAILSSIIISAVL 489
Qy 461 GLIDYEAAILHFKVDKDFVVCMSAYIGVVGSGVEIGLVIATVISVLRVLFTIARPTFV 520
Db 490 GLIDYESAYITWIKVDLDFLACMGAFGEVFSSEYGLLIAVAISLAKILLQVTRPTVL 549
Qy 521 LGNIPNSVIRNVEHYQNAKHVPGMLTLEIDAPIYFANASYLRIRIRWIDEEERIKAT 580
Db 550 LGNLPRTIIRNVEQYDPAKYFQVLIVRDSALYFNSVYKVERILRILWRIDEEEOQDQ 609
Qy 581 GETSLQVVIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNKSKPON 640
Db 610 KLTKTEPLIVDLSPVIDIDTSGIHAELEALAKLEKRIQLVLTNPGPAVIQKLSAKFTD 669
Qy 641 HLGKKWYLYTVEAV 655
Db 670 MIGEDNIFLTVGDAV 684

RESULT 4
US-10-424-599-207945
; Sequence 207945, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 207945
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(659)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29800C.1.pep
US-10-424-599-207945

Query Match 53.2%; Score 1850; DB 15; Length 659;
Best Local Similarity 55.9%; Pred. No. 3.8e-151;
Matches 359; Conservative 122; Mismatches 153; Indels 8; Gaps 4;

Qy 41 VHQVEVPPPPFFKSLKYSKLTFFPDPLRQPKNKPKASKKFMGLQFFPFIFEWAPKYT 100
Db 20 VHQVPPPHKSTLQKLGRLEKTFPPDPLRQPKGQPLKRLILGAYVPPILQWGPKN 79
Qy 101 FQFLKADLIAGITIASLAIPQGISYAKLANLPITGLYSSFIPIPLIYAMGSSRDLAGVT 160
Db 80 LKLFKSLVSLGSLTASLDIPQGSYAKLASLPILGLYSSFVPPVAVILGSSKDLAVGP 139
Qy 161 VAVGSLMGSLMNAVDNPDNPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLGHATI 220
Db 140 VSTASLVGSLMNAHQEVSPTTDPILFLQALFTSTLFAGLFOALLGILRLGLIIFLISKAIL 199
Qy 221 IGFMGGAATVVCVQLKLSILGLEHFTHGADIISVMRSVFTQTHWRWESAVLCGVFIFFL 280
Db 200 IGFMGAAATVSVLQKLSILGLIHTFTNQMGILPMTSVFHNIEHWSQTLIMGICFLVLL 259
Qy 281 LSTRYSKKEPRFVWSAMAPLTSVILGSLIVFTHAEKHGVEVIGELKGLNPPSLNL 340
```

```
Db 260 LLARHVSIRKPKLFWVSAGAPLMCVIISTLLVPAIKAQNHGISAIGKLOQGINPPSWNML 319
Qy 341 VFVSPYMTAVTKTIGIVVGIISL-AEGIAVGRSPAMTKYNNIDGNKEMIAIGTNNVVGSGFT 399
Db 320 LPHGSHLGLVMKLTGLITLSTXEGIAVGRTEFAALKYKNYKVDGNKEMMAIGFNNVVGSGFT 379
Qy 400 SCVLTGTPFSRANYNAGCKTAASNIIMSLAVMLTLLFLPLFHYTPLVVLISAILVSAM 459
Db 380 SCVYTGAFSRSANVNAAGAKTAVSNVMSVTVMTLLFLMPLFQYTPNVVLGAIIVTAV 439
Qy 460 LGLIDYEAAILHFKVDKDFVVCMSAYIGVVGSGVEIGLVIATVISVLRVLFTIARPTF 519
Db 440 IGLIDLPACANIWKIDKDFVVMWMTAFGLVLFISVOGLALAVGLSTLKLLOITRPTKV 499
Qy 520 VGNIPNSVIRNVEHYQNAKHVPGMLTLEIDAPIYFANASYLRIRIRWIDEEERIKAK 579
Db 500 MLGKPGTDIYRNLDQYKEAVRIPGFLIISIEAPINFANITYLNERTLWIEEEDNIKE 559
Qy 580 TGSETSLOVVIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNKSKPQ 639
Db 560 --QLSLRFLVLEMSAVSVDTSIGLSLFLKATLEKKGVLELVNPLAEVIEKLKKADEA 617
Qy 640 N-HLGKKWYLYTVEEAVGACNPNLRASKTNPKKDETEGWNV 680
Db 618 NDFIRADNLFITVGEAVA-----SLSAMKQSGSTITEGAHTI 655

RESULT 5
US-10-437-963-198336
; Sequence 198336, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198336
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94005C.1.pep
US-10-437-963-198336

Query Match 52.1%; Score 1811.5; DB 16; Length 666;
Best Local Similarity 53.8%; Pred. No. 8.3e-148;
Matches 340; Conservative 122; Mismatches 167; Indels 3; Gaps 3;

Qy 43 QVEVPPPPFFKSLKYSKLTFFPDPLRQPKNKPKASKKFMGLQFFPFIFEWAPKYTQ 102
Db 36 EVNLSGRPPFAEKLSWSDLAETFPDPPFGFGALPPARRAWCAKTFVPALDWPVRLGD 95
Qy 103 FLKADLIAGITIASLAIPQGISYAKLANLPITGLYSSFIPIPLIYAMGSSRDLAGVTVA 162
Db 96 KFKFDLLAGITIASLAIPQGISYARLANLPITGLYSSFPVPLMYAVFGSSNNLAVGTVA 155
Qy 163 VGSLLMGSLMNAVDNPDNPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLGHATIIG 222
Db 156 AASLLIASIETEAADENPOLYQLFYTAAPFTGLFQALGVFRGLIVDFLSRSTIG 215
Qy 223 FMCGAATVVCVQLKLSILGLEHFTHGADIISVMRSVFTQTHWRWESAVLCGVFIFFLS 282
Db 216 FMGGTAMIIILQOFKGLLGKHKHFTTKTDIISVLHSTYHWRHKKWQSAVLGICFLFLMS 275
```

```
QY 283 TRYFSKRRPRFFVWSAMAPLTSVILGSLLYVFTHAEKHGVEVIGELKKGILNPPSLTNLVF 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 276 SKHLKKLPKLFVWSALAPFMVVIGGIPAPLVKGDHGPVIGNLKNGINPLISIAQLTF 335

QY 343 VSPYMTAVKTVIGVIGIISLAEGIAVGRSPAMYKNYINDGNKEMIAIGTMNVVGSFTSCY 402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 336 QSRHMKTAVKAGLLSGILALAEGLAVGRSLAMVKNQINDGNKEMIAFGIMNIAGSFTSCV 395

QY 403 LTTGPPRSRNVNACKTAASNIIMSLAVMLTLFLPLFHYTPLVLSIAIIVSAMGL 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 396 LTTGPPFSKSAVNHAGCKTPMSNVMSVIMLVLLFLAPLFKYTPLVASLSIIIVAMIGL 455

QY 463 IDYEAAILHLPKVDKDFVFCMSAYIGVFGSVIGIIVIAIVISVLRLVLIARPRTFVLG 522
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 456 VKYKECHLYKVDKDFCICWAFLVGVFTWVAGLSASVGLSVRALLYVARPATCKLG 515

QY 523 NIPNSVIYRNVEHYQNAKHVPQMLILEIDAPIYFANASYLRERITWIDEEERIATGE 582
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 516 NIAGSETFRDVQYPOAKSIPGLVLQLGSPYFVNAGYLRERILRWEDENLCKSVGH 575

QY 583 TSLQYVLIOMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNKSKF-QNH 641
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 576 -DLQYLVLDLGGVTSVDNSGVGMLELVHKSLERRGITIALNPRLEVTEKVLVSGYRDI 634

QY 642 LGKKIYLVTVEEAVGACNFMNRASKTNPKKDE 673
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 635 LGDEWFLVTKDAITACRYALQISR-NKGEDE 665

RESULT 6
US-10-437-963-179488
; Sequence 179488, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179488
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76946C.1.pep
US-10-437-963-179488
```

```
Query Match 51.0%; Score 1773.5; DB 16; Length 662;
Best Local Similarity 52.6%; Pred. No. 1.6e-144;
Matches 328; Conservative 128; Mismatches 166; Indels 1; Gaps 1;

QY 34 GNMNFRVHVQVPPPPQPFKSLKYSLKETFFPDPLRQFNKPKASKKFMGLGQFFPPIF 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 29 GHHHHHHGHKVEPPPKLLIDEFTDAVKETFFADDPRLRQYKQPKMSKKVLSIQNFFPVL 88

QY 94 EWAPKTFQFLKADLIAGTIIASLAIPQGISYAKLANLPPIGLYSSFPPIIYAMGSS 153
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 89 DMRGHTFRKRGDLVSGLTIASLCIPQDIGYAKLAGLLPNYGLSYSSFPPIIYAMGSS 148

QY 154 RDLAVGTAVGSLMGMSLSNADVPNEDPKLYHLAFTATLTFAGVFOAALGLFRGLGLVD 213
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 RDLAIGFVAVSLLGLTILQNEEDPKKQNEEYTRLAFTATTFAGVTQAVLGLFRGLGL 208
```

```
QY 214 FLSHATIIIFMGGAATVVCLOQLKSLGLEHFTHGADIISVMRSVFTQTHE-WRWSAYL 272
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 FLSHAIVGFMAGAAITIALQQLKGLGIANFTKTDIIISVMKSVGNVHGMNQTIILI 268

QY 273 GCVTFIFLLSTRYFKRPRFFVWSAMAPLTSVILGSLLYVFTHAEKHGVEVIGELKKG 332
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 269 GASFLAFLVAKYIAKKNKLEFWAAIPLTSVIIITLFFVYITRADKHGVIVKIKGI 328

QY 333 NPSPITNLVVFVSPYMTTAVKTVIGVIGIISLAEGIAVGRSPAMYKNYINDGNKEMIAIGTM 392
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 329 NPSPASLIYFSGPNLMKGFRIAGIAGITLTAIAGRTFAGLKDKYIDGNKEMWALGTM 388

QY 393 NVVGSFTSCYLTGTPFSRSVAVNAGCKTAASNIIMSLAVMLTLFLPLFHYTPLVLS 452
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 389 NIVGSMTSCYVATGSFSSRNVNAGCQATVSNVMSIVLLTLELITFLPKYTPNAILS 448

QY 453 AIIVSAMGLIIDEYEAAILHLPKVDKDFVVCMSAYIGVFGSVIGIIVIAIVISVLRLV 512
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 449 SIIISAVLGLVDYHTAYLIWKVDKDLACLGAFFGVFSSEYVEGLIIAVALSLAKILQ 508

QY 513 IARPTFVLGNTPNSVIVRNVEHYQNAKHVPQMLILEIDAPIYFANASYLRERITRWIDE 572
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 509 VTRPRTVLGNUPRTILYRNIDQYPEARLVPGVIVRVDLSAIYFTNSNVKRIILRWLRD 568

QY 573 EERIKAETSLQYVLIOMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKK 632
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 569 EEERQEQKLOKTEFLIVELSPVIDITSGIHALEDLPRALEKRIQLILANPGPAVILK 628

QY 633 LNKSQFQNLGKKKIYLVTVEEAV 655
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 629 LRSAKFTDLIGEDKIFLTVGDV 651

RESULT 7
US-10-425-115-260426
; Sequence 260426, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 260426
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_169123C.1.pep
US-10-425-115-260426
```

```
Query Match 50.7%; Score 1762; DB 16; Length 681;
Best Local Similarity 53.9%; Pred. No. 1.7e-143;
Matches 334; Conservative 120; Mismatches 162; Indels 4; Gaps 2;

QY 41 VHQVPPPPQPFKSLKYSLKETFFPDPLRQFNKPKASKKFMGLGQFFPPIFEWAPKYT 100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 LHKVSLPERRTTAKALRQRLAEVFFPDPLHFKQKQSSARRLVLAHYVFFPIFQWGSAYS 118

QY 101 FOFLKADLIAGTIIASLAIPQGISYAKLANLPPIGLYSSFPPIIYAMGSSRLAVCT 160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 PRLRSDLVAGTIIASLAIPQGISYAKLANLPPIVGLYSSFPPIIYALLGSSRLAVGP 178

QY 161 VAVGSLLMGMSLSNADVPNEDPKLYHLAFTATLTFAGVFOAALGLFRGLGLIVDFLSHAT 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 VSIASLVWMSLMRDVSPDEQLLLQLAFTATTFAGVFOASLGLFRGLGLIVDFLSKATL 238

QY 221 IGFMGAATVVCLOQLKSLGLEHFTHGADIISVMRSVFTQTHEWRWSAVLGCVFIFPL 280
```

```

Db      239  TGFMGGAIVVSQQLKGLGISHFTSHMGFLDVMRSVVRHDEWKQTIWMSAFILAIL 298
Qy      281  LSTRYFSKGRPRFPFWVSAMAPLTSVILGSLVYFTHAEKRGVEVIGELKKGLNPPSLTNL 340
Db      299  LLTRQISARNPKLFWVSAGAPLASVIISTILSPFW--KSPSISVIGILPRGVNPPSANML 356
Qy      341  VFVSPYMTTAVKTIIVVGIISLAEGIAVGRSPAMYKYNIDGNKEMIAIGTMMVVGSGFTS 400
Db      357  SFGSGYVALTIKTGIMTGILSTEGIAVGRTPFASINNYQVDGNKEMMAIGLMMAGSCAS 416
Qy      401  CYLTGPFSSAVNAGCKTAASNIIMSLAVMLTLLFPLPHYTPLVLSAIIYSAML 460
Db      417  CYVTGSGSSAVNYSAGCKTALSNVMAAALVTLFLMPLPHYTPNVLSAIIITAVV 476
Qy      461  GLIDYEAAHLFKVDKDFVVCMSAYIGVVGVSVEIGLVAIIVISLVRLLFIARPRTFV 520
Db      477  GLVDVRGAARLWKVDKDLFLACVAAPLGLVLSVQTGLGVAVGLSLFKVLLQVTRPNVVV 536
Qy      521  LGNIPNSVIYRNVEHYONAKHVPQMLILEIDAPIYPANASYLRERITRWIDEEBERIKAT 580
Db      537  EGLVPGTQSYRSVAQYREAVRVPGLVVGVSAYIFANSMLYVERVMRYLRDEERALKS 596
Qy      581  GETSLQYVIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKLNKSKFQN 640
Db      597  NHPISIRCVLMDGAAVIDTSGLDALSELKKVLDKRNIELVLNPGVSVAEAFNPSAVGE 656
Qy      641  HLGKRWIYLTVEEAV--GAC 658
Db      657  SFGSGRLFFPSVAEAVAGAC 676

RESULT 8
US-10-437-963-175893
; Sequence 175893, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 175893
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73694C.1.pap
US-10-437-963-175893

Query Match 49.98; Score 1733.5; DB 16; Length 656;
Best Local Similarity 52.98; Pred. No. 4.7e-141;
Matches 333; Conservative 113; Mismatches 162; Indels 21; Gaps 3;

Qy      41  VHOVEVPPPPFFKSLKYSLKETFFDDPLRQFNKPASKKFMGLGQFFPIPEWAPKYT 100
Db      49  LHKVSVPERSTAKALRQLAEVFFDDPLHQFNQSSARRLVLAQYFFPIHMGSDYS 108
Qy     101  FQFLKADLIAGITIASLAIPQGISYAKLANLPILGLYSFIPPLIYAMGSSRDLAGVT 160
Db     109  LRLLRSD-----GISYAKLANLPILGLYSFVPPPLIYSLIGSSRDLAGVP 154
Qy     161  VAVGSLMGSMLSNAVDPNEDPKYLHLHATLTFAGVFOAALGLFRLGLIYDFLSHATI 220

```

```

Db      155  VSTASLWMSMLRQAVSPDQEPILYLQLAFTSTFFAGVQFQASLGFLRGLFIVDFLSKATL 214
Qy      221  IGFMGGAATVVCVLQQLKSLGLEHFTHGADIISVMRSVFTQTHERWESAVLGCVFIFFL 280
Db      215  TGFMGGAATVVCVLQQLKSLGLEHFTHGADIISVMRSVFTQTHERWESAVLGCVFIFFL 274
Qy      281  LSTRYFSKGRPRFPFWVSAMAPLTSVILGSLVYFTHAEKRGVEVIGELKKGLNPPSLTNL 340
Db      275  LLTRHISARNPKLFWVSAAAPLTSVIISTISFVSKA--HGISVIGDLPLKGLNPPSANML 332
Qy      341  VFVSPYMTTAVKTIIVVGIISLAEGIAVGRSPAMYKYNIDGNKEMIAIGTMMVVGSGFTS 400
Db      333  TFGSGVGLALNTGIMTGILSTEGIAVGRTPFASINNYQVDGNKEMMAIGVMMAGSCAS 392
Qy      401  CYLTGPFSSAVNAGCKTAASNIIMSLAVMLTLLFPLPHYTPLVLSAIIYSAML 460
Db      393  CYVTGSGSSAVNYSAGCKTASVNIWMAASVLAIVTLFLMPLPHYTPNVLSAIIITAVI 452
Qy      461  GLIDYEAAHLFKVDKDFVVCMSAYIGVVGVSVEIGLVAIIVISLVRLLFIARPRTFV 520
Db      453  GLIDVRGAARLWKVDKDLFLACMAAFGLVLSVQVGLAIVAGISLFLKILLQVTRPNVVV 512
Qy      521  LGNIPNSVIYRNVEHYONAKHVPQMLILEIDAPIYPANASYLRERITRWIDEEBERIKAT 580
Db      513  KGVPFGTASYSRWAQYREAVRVPGLVVGVSAYIFANSMLYGERIMRFLREEDERAAC 572
Qy      581  GETSLQYVIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKLNKSKFQN 640
Db      573  NQCPVRCIILDSAVAAIDTSGLDALAEKLVLEKRNIELVLNPGVSVTERLYNSVVGK 632
Qy      641  HLGKRWIYLTVEEAVGACNPNLRASKTNP 669
Db      633  TFGSDRVFPFSAEAVAAA-----PHKTOP 656

RESULT 9
US-10-424-599-273959
; Sequence 273959, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 273959
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(397)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89405C.1.pap
US-10-424-599-273959

Query Match 49.6%; Score 1724; DB 15; Length 397;
Best Local Similarity 81.0%; Pred. No. 1.5e-140;
Matches 342; Conservative 19; Mismatches 29; Indels 32; Gaps 3;

Qy      258  VFTQTHEWESAVLGCVFIFLLSTRYFSKGRPRFPFWVSAMAPLTSVILGSLLVYFTHA 317
Db      7  VFSQTHEWESAVLGCVFIFLLSTRYFSKGRPRFPFWVSAMAPLTSVILGSLLVYFTHA 66
Qy      318  EXHGVGVIGELKKGLNPPSLTNLNVFVSPYMTTAVKTIIVVGIISLAEGIAVGRSPAMYKN 377
Db      67  EXHGVGVIGELKKGLNPPSATDLVFPVSPYNGTAIKTGTGTGTTIALAEGIAVGRSPAMFKN 126

```

```
QY 378 YNIDGNKEMIAIGTMNVVSGFTSCYLTTPGPFSSAVNNAGCKTAASNIIMSLAVMLTL 437
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 127 YHIDGNKEMIAIGTMNIFSGFTSCYLTTPGPFSSAVNNAGCKTAASNIIMAVMLTLL 186

QY 438 FLTPLFHYTPLVLSAIIIVSAMGLIDYEAAILFKVDKDFVVCMSAYIGVFGSVEIG 497
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 187 FLTPLFHYTPLVLSAIIIVSAMGLIDYEAAILFKVDKDFVVCMSAYIGVFGSVEIG 246

QY 498 LVTAIVISLVRLLFTARPRTEVLGNIPNSVIYRNVEHYQNAKHVPCMLILEIDAPIYA 557
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 247 LVTAIVISLVRLLFTARPRTEVLGNIPNSVIYRNVEHYQNAKHVPCMLILEIDAPIYA 305

QY 558 NASYLREIRTRWIDEEBERIKATGETSLOXYIIMDSAVGNIDTSGISMLEEVEKKITERRE 617
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 306 NASYLREIRTRWIDEEBERIKA-----GMFDSXACMLGEI----- 341

QY 618 LQVLVNPVSEVMKLNKSKFQHLKKWYLYTVEEAVGACNPNLRAKTNPKKDBTEGW 677
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 342 -----PRSEVMKLNKSKFLDELQKWYLYTVEEAVGACNPNLHSHYKPNPMKDESEG 394

QY 678 NN 679
||
Db 395 NN 396
```

```
RESULT 10
US-10-437-963-179495
; Sequence 179495, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179495
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(653)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76952C.1.pep
US-10-437-963-179495
```

```
Query Match 49.6%; Score 1723.5; DB 16; Length 653;
Best Local Similarity 53.2%; Pred. No. 3.4e-140;
Matches 328; Conservative 119; Mismatches 165; Indels 5; Gaps 2;

QY 42 HQVEVPPPPQPFKSKLSKLTETFPDDPLRQFNKPKASKFMLGLQFFPFPIFEWAPKTYF 101
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 28 YNVAAPPKKNLLAEFAAGTAKETFFSDPMRYKQDQPRSRKMLALQHVFPVFWGQYTL 87

QY 102 QFLKADLIAGITIASLAIPQGISYAKLANLPPIILGYSSFPPIIYAMGSSRDLAGVTV 161
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 88 AKFKGLDIAAGLTASLVIPODICYAKLANLPPIIGHSSFPVPLIYALMGTSRELAMGV 147

QY 162 AVGSLLMGMSLNAVDPNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLSHATII 221
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 148 AVISLLLTLLQOEIDSKNPLDYRLAFTATLFAGVFOAALGLFRLGLIVDFLSHATII 207
```

```
QY 222 GFMGGAATVVCLOQLKSLIIGLSEHFTHGADIISVMRSVFTQTH---EWRWESAVLCGVFIF 278
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 208 GFMAGAAITIALQQLKGLPLGIANFTKTKTDIISVMKSVGCVNVHGXQNNWQTILIGASFLA 267

QY 279 FLILSTRYFSKKPRPFWSAMAPLTSVLGSLLYVFTTHAEKKGVEVIGELKGLNPPSLT 338
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 268 FLILVA--XGKQNKCLFWWPAIAPLISVIITSTLFTVITRADKQGVAVIKVKNKGINPPSAS 325

QY 339 NLVFSVPYMTTAVKGTIVVGIISLAEGIAVGSFAMYNKYNIDGNKEMIAIGTMNVVGSF 398
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 326 LIFFTPGYLLKFKFGLVAGMISLTEAIAVGRFTAGLNDYQIDGNKEMIALGTMNVVGS 385

QY 399 TSCYLTGTPFSSAVNNAGCKTAASNIIMSLAVMLTLLFLTPLFHYTPLVLSAIIVSA 458
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 386 TSCYIATGFGFARSANVMAGGKTPMSNIVMSTVLLALLMITPLPKYTENATISSIIISA 445

QY 459 MGLIDYEAAILFKVDKDFVVCMSAYIGVFGSVEIGLVIATVISLVRLLFTARPR 518
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 446 VLGLFDFESAYLIWKVDKLDLDFMACLGAFGLGVIFSSVEYGLLIAVVISLIKVLHVT 505

QY 519 FVLGNIPNSVIYRNVEHYQNAKHVPCMLILEIDAPIYANASYLREIRTRWIDEEBERIK 578
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 506 ALLGNLPRITIIYRNVEQYPTATKVPGLMILVRVDSAIYFTNSYVYKERMILWRDDEEHOK 565

QY 579 ATGETSLOXYIIMDSAVGNIDTSGISMLEEVEKKITERRELQVLVNPVSEVMKLNKSKF 638
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 566 EQKLKPIELIVDLSPVNDIDTSGIHAPKELLRTLEKQIQLIFANPGAAVIQKLSAKF 625

QY 639 QNHGKKNWYLYTVEEAV 655
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 626 TELIGEKEICTVGDV 642

RESULT 11
US-10-425-115-366870
; Sequence 366870, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 366870
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(662)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_97750C.1.pep
US-10-425-115-366870

Query Match 49.6%; Score 1722.5; DB 16; Length 662;
Best Local Similarity 53.9%; Pred. No. 4.3e-140;
Matches 343; Conservative 110; Mismatches 168; Indels 15; Gaps 6;

QY 41 HQVEVPPPPQPFKSKLSKLTETFPDDPLRQFNKPKASKFMLGLQFFPFPIFEWAPKTY 100
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 35 VHKVAPPARSTASKKVRVKTETFPDDPFRAFQKQPPGTQMLMAVRYLFPILDWVPSTYS 94

QY 101 FOFLKADLIAGITIASLAIPQGISYAKLANLPPIILGYSSFPPIIYAMGSSRDLAGVTV 160
|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 95 LSLFSLDVLAVGTLIASLAIPQGISYAKLASUPPIIGLYSSFPVPMVYAVLGSRRDLAVGP 154

QY 161 VAVGSLLMGMSLSNA-VDPNED----PKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFL 215
```

Db 155 VSISLWGPCCXAGREPHRGADAVPAAGUH-----XTLFGVLQASIGILRLGFGVIDFL 210  
Qy 216 SHATIIIFGMAATVVCLOQLKSLGLEHETHGADIISVNRSVFTQTHWRWESAVLGCY 275  
Db 211 SKATLVGFMAAGATIAVALQALKALGIVHFTTENGIVPVMAVVFHHTSEWSQTIILGVC 270  
Qy 276 FIFELLSTRFSGKRPFFWWSAMAPLTSVLGSLLYVFFTHAEKHGVEVIGELKGLNPP 335  
Db 271 FLVFLLSARHVSIRWPKLFWSCAPLASVTISTLLVFLFKAQNHGSIIGQLKCLNRP 330  
Qy 336 SLTNLWVSPVMTTAVKTIIVGVIISLAEGIAVGRSFAMYNIDGNKEMIALGTMMNV 395  
Db 331 SWDKLLFDYALGHTMTKGLVTLGIIISLATEGIAGRTFASLKDQIDGNKEMIALGMMNV 390  
Qy 396 GSFTSCYLTTPGFSRSVAVNAGCKTAASNIISLAVMLTLLFPLFHYTPTLVLSAII 455  
Db 391 GSCTSCYVTTGAFSRSVAVNAGCKTAMSNVIMALTVMVTLFLMPLFVYTPNVVLGAI 450  
Qy 456 VSAMGLIDYEAAIHLFKVDKDFVWCMSAYIGVWFGSVEIGLVIIVSLRVLLFIAR 515  
Db 451 IAAVIGLIDFPVAVHIWMDKDFLVCVCAFAGVIFISVQEGIAIAVGISIFRLMQITR 510  
Qy 516 PRFVLGNIPNSVIRNVEHYQNAKHVGMMLILEIDAPIYFANASYLRERITRWIDEE 575  
Db 511 PKMVQGNIGTDIYRDLHYKQAVSGFLIILAIEAPINFANSYLNRIKRWI--EE 568  
Qy 576 RIKATGETSLQYIIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNK 635  
Db 569 SFEQDKHTLHFIILDSAVPAIDTSGIAFLIDIKKSIKRGLELVLVNPTGEXHENTT 628  
Qy 636 SK--FQNLGKKKIYLTVEAVGACNPNLRASKTNP 669  
Db 629 CKRGXKTIILGACIXTTGEAI--ASLSALAKMTKP 662

RESULT 12

US-10-437-963-174545  
; Sequence 174545, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 174545  
; LENGTH: 659  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_72475C.1.pep  
US-10-437-963-174545

Query Match 45.0%; Score 1563.5; DB 16; Length 659;  
Best Local Similarity 47.7%; Pred. No. 2.6e-126;  
Matches 299; Conservative 121; Mismatches 194; Indels 13; Gaps 3;  
Qy 50 OPPEFKSLKYSKLTFFDDPLRQFNKPKASKKFMGLQFPFFIFEWAPKTYTFOFLKADLI 109  
Db 29 RPFCAARAALKDTLFPDDPFRGLGCMPPARRAWRVFVFPALDMGAGYSAASFWYDLL 88  
Qy 110 AGTITIASLAIPOGISYAKLANLPILGLYSFPTPLIYAMGSSRDIAVGTVAAGSLIMG 169

Db 89 AGVTIASLAIPOGISYATLAGIPPPVIGLYSCFVPPLYVAVMGSSRNGLVGPVATSSLLVA 148  
Qy 170 SMLSNAVDNEDPKLYLHLAFTATLTFAGVQAAALGFLRLGLIYDFLSHATIIFGMAAT 229  
Db 149 SIUGGKVRASDDQQLYTLQVFTSAFPTGLVLAALGLRLGLILVDFMSRPAITFGMGTAI 208  
Qy 230 VVCLQOLKSLGLEHETHGADIISVNRSVFTQTHWRWESAVLGCYFIFELLSTRFSEKK 289  
Db 209 VIMLOQLKGLGHTFTTKTIDIVSLRYIFHNTHOMQOSTVLGVCFILFLVFTQVRRR 268  
Qy 290 RPRFFWWSAMAPLTSVLGSLLYVFFTHAEKHGVEVIGELKGLNPNPSLTNLVFPVPMYT 349  
Db 269 RPKLFWWSAMSPLLVVVGCVPFLIKHKGPIVGTLKRGINPSSISOLKQPEYGV 328  
Qy 350 AVKTIVGVIIISLAEGIAVGRSFAMYNIDGNKEMIALGTNNVVGVSFTSCYLTTPGFS 409  
Db 329 AMKAGFVSGMLALAEVAVGRSFAAMKKERIDGNKEMVAFGLMNLIGSFTSCYITGAFS 388  
Qy 410 RSNVYNAAGCKTAASNIISLAVMLTLLFPLFHYTPTLVLSAIIIVSAMGLIDYEAAI 469  
Db 389 KTAENVHAGCRTAMSNVMSVCMALVLVALPLFRHTPLVALAAIITSSMLGLVKGREIR 448  
Qy 470 HLPKVDKDFVWCMSAYIGVWFGSVEIGLVIIVSLRVLLFIARPRTFVLGNI----- 524  
Db 449 RLYEVDKADFVCAALGLGVVFTMTGLGVAIVSLRALLHVARPSTSKLGRVSCGS 508  
Qy 525 ----PNSVIRNVEHYQNAKHVGMMLILEI-DAPIYFANASYLRERITRWIDEEERIKA 579  
Db 509 AGAADDHAFCDVAQYPGAATAPSLVLQVAGSPVCFANAEYLRERIARWVEDEE--KA 565  
Qy 580 TGSETSLOQYIIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLNKSKFQ 639  
Db 566 VAGEDLLYVLDIGGVTAIDSPGIEMLREVHGLERKGMKMAVTNPRMAVEKLVLGSLA 625  
Qy 640 NHLGKKKIYLTVEAVGACNPNLRASK 666  
Db 626 ELVGESWFMFLSNGDALLAACRYTLQSK 652

RESULT 13

US-10-424-599-195017  
; Sequence 195017, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 195017  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_18127C.1.pep  
US-10-424-599-195017

Query Match 43.2%; Score 1500; DB 15; Length 653;  
Best Local Similarity 48.0%; Pred. No. 8.1e-121;  
Matches 297; Conservative 126; Mismatches 178; Indels 18; Gaps 7;  
Qy 48 PPQPFKSLKYSKLTFFDDPLRQFNK-----PASKKFMGLQFPFFIFEWAPKTYT 100  
Db 38 PPSF-WRVVADSVSKT-----ISHYKHLSSLIDQPCCTLLSLVQLVFPFILAAGRNT 90  
Qy 101 FQFLKADLITAGITIASLAIPOGISYAKLANLPILGLYSFPTPLIYAMGSSRDIAVGT 160  
Db 91 ATKFRKOLLAGLTIASLCIPQSIGYATLAHLDPOYGLTYSVPPPLIYAVMGTSREIAIGP 150

Qy		161	VAVGSLMGSMLSNAVDPNEDPKLJHLHAFATLATLPAGVFOAALGLFRGLIVDFLSHATI	220
Dd		151	VAVVSLLSSMMEKLVDPATDPVGVTUKILLATLPFAGIFOTSFGILRLGFLVDFLSHAAI	210
Qy		221	IGFWGGAAVVCLQQKSIILGLEHEFTHCADIIISVMRSVFTOTHE-RRWESAVLGCVPIFF	279
Dd		211	VGFVAGAAIIVGLQQKGELGTHTFTTKTDIVSVMKAVEAVHNPPNPNFLGCSELVF	270
Qy		280	LLSTRYFSKKRPRFPWFWSAMAPLTSVIUGSLLVYFTHAEKHGEVEIGELKKGLNPSPSUTN	339
Dd		271	ILTTRFLGKRKKLFWLASISPLSVSVLTLIIVFLTRADKNGVKIVKHVKGLNPSLSHQ	330
Qy		340	LVFVSPYMTTAKVTGIIVGIIISLBAGIAVGRSFAMYKNYNIDGNKEMTAIGHMVNVGSGTF	399
Dd		331	LDFNNPVI GEYAKI GLVVAUVALTESI AVGRSFASIKGYQLDGNKMMSIGLTNIIGSTF	390
Qy		400	SCYLTTGPFSFSAVNYNAGCKTAASNITMSLAVMLTLPLTPLFHYTPLVLVLSAIVSAM	459
Dd		391	SCIYATGSFSTRAYNAAAGCELTLSNINIMATLVLSLOFLTLLTYTPTAILASVILSAL	450
Qy		460	LGLIDYEAAIHLFKVDKPDFVVCMSAYIGVGVGFSVEIGLVAIVISLVRLLFIARPTF	519
Dd		451	PGLIDLSEAYKIWKVKDKIDFLACAGAFFGVLFAFASVEIGLLVAVVISFSKILIISRPGTE	510
Qy		520	VLGNIPNSVINRVNEHYQNAXHPGMLEI-DAPIYFANASYLBERITRWIDESE-BRI	577
Dd		511	TLGKIPGTDLCFDVQYPMVAVKVPQVMIRVKSALLCFANANFVERIIRKWVTEESEDD	570
Qy		578	KATGETSLQVYIIDMSAVGNIDTSGISMLEVEKKITERRELQQLVLVNPVSEVMKCLKNSK	637
Dd		571	KGNERSITQLVILDTSNLVNIDTSGITALALEHLKLSLSSQGKQLATANPRQVITHKLKYSN	630
Qy		638	FQNHGKKKIIVLTVEEAVG	656
Dd		631	FVGKIGGR-VELTVEEAVG	648

```

RESULT 14
US-10-425-114-69272
; Sequence 69272, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69272
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMST02400016G05_FLI.pep
US-10-425-114-69272

```

	Query Match	42.0%	Score 1460;	DB 15;	Length 575;
	Best Local Similarity	49.3%;	Pred.No. 2e-117;		
	Matches 281; Conservative	120;	Mismatches 165;	Indels 4;	Gaps 4;
Qy	90 PPIFEWAPKYTEFQLFKADLLAGITIASLAI PQGISYAKLANLPILGLYSSTFPPLIYAV	149	:	:	:
Dd	2 FPIAWGNVTATWKFDKLTLA GLTSLCIPQSIGYATLAHLDPQGYLGIVTSVPPLIYV	61	:	:	:
Qy	150 MGSSRDLAGVTVA VGS LLMGSMLSNAVDPNEDPKLYLH LAFTATLT PFAGVFQAALGFLRLG	209	:	:	:

62 MGTSEINAIQVAVVSVLLSSMWEKLVDPADTPVGYTKLILLATLIFAGITFQTSFGLLRUG 121  
 210 LIVDFLSHATIIIGFMGGAATVVCLOQLKSIIIGLEHFTHGADIISVMRSVFTOTHE- WRWE 268  
 122 FLVDFLSHAAIVGVFAGAAIIVIGLQQLKGLIGITHTFTTKTDIVSVNKAWEAVHNPWR 181  
 269 SAVLGCVFIFILLSTRYFSKKRPFVWSAMAPLTSVILGSLIVVYPTHAEKHGVEVIGEL 328  
 182 NFILGCSFLVILFTRCLGKRKKLFWLASISPLSVSVVSTUIVFTITRADKNGKIVKHV 241  
 329 KKGILNPPSLTNLNVFVSPYMTAVTKTIGVVGIIISLAEIGVGRSFAMYNKYNIDGNKEMIA 388  
 242 KGLNPPSIHOLDENNPIGEVAKIGLVAVVALTESAVGRSFASIKGYQLDGNKEMWS 301  
 389 IGTWNVGFSFSCYLTTFGFRSNAVYNAGCKTAASNIIMSLAVMLTLFLTPLFHYTPL 448  
 302 IGLTNIIIGSFSCVATGFSRFTAVNYAAGCETLVSNIVMAITVLIISQLTKLJYTP 361  
 449 VVLSAIVSMGLGIDYEAATHLKFVKDFDVVCMASVIGVVGVSVEIGLVIATVSVLR 508  
 362 AILASVILSALUGLIDLSEAYKIMWKDKIDFLACAGAFGVLFASVEIGLVAVVISFSK 421  
 509 VLLFIARPTFVLGNIPNSVIRYVHEHYQNAKHVFGMLILEI-DAPIYFANASYLRERIT 567  
 422 IILSIRPGTETLGLKPGTDLFCDDVYQVPMVAVKVPGVMIIRVKSALLCFANANFVRERI 481  
 568 RWIDEEE-ERIKATGETSLOYVIDMSAVGNIDTSGISMLERVKKITERRELVLVNPV 626  
 482 KWTVEEEDDKGNSRSTIQFLDLDTSLNVIDTAGITALEBLHKLSLSHGHQOLAIPNR 541  
 627 SEVMKLNKSKFQNHGLKKWILYLTVEEAAG 656  
 542 WQVIHLKLVSNFVGKIRG-VFLTVEEAAG 570  
 RESULT 15  
 US-10-425-115-228851  
 ; Sequence 228851, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kowalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 228851  
 ; LENGTH: 405  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MET4577\_140304C.1.pep  
 US-10-425-115-228851

	Query Match	41.3%	Score 1436;	DB 16;	Length 405;
	Best Local Similarity	67.3%;	Pred. No. 1.5e-115;		
	Matches 268;	Conservative 65;	Mismatches 61;	Indels 4;	Gaps 1
Qy	275 VFIFELLSTRVFSKRRPRFFWVSAAPLTSTVILGSLIVYFTHAEKHGVETVELKKGLNP	334	:	:	:
Dd	3 IICFFFIQ----	SKRPRKLFWISAAAPLTSVLGSGVLVYLTHAENHGEVIYGLKKGLNP	58	:	:
Qy	335 PSLTNLVFVSPMYMTAVKTGIWGIIISAEGIAVGSRPFAMYNKNYINDGNKEMIAIGTMNV	394	:	:	:
Dd	59 PSVTSLQPSPPYMMALKTGIIITGVIALAEGIAVGRSFPAMFNHYIDGNKEMIAIGTMNV	118	:	:	:
Qy	395 VGSFTSCYLTTGPFRSAVNNAACKTAASNIIMSLAVMLTLFLTPTLFHYTPTPLVLSAI	454	:	:	:
Dd	119 LGSITSCLYTGGPFRSAVNNAACRTAMSNVNSLAVMVLTLLFLTPHFHTPTPLVLSAI	178	:	:	:



Qy	455	IVSAMGLIDYERAIHLPKYDKDFVVCMSAYIGVWFGSVGEIGLVIAIVISVLRLVLLFIA	514
Db	179	IVSAMGLVDGAAHLWVRVDKDFVCAGAYLGWFGSVGEGLWAVASLLRLVLRFGA	238
Qy	515	RPRTFVLGNIPNSVIYRNVEHYONAKHVPGMILEIDAPIYFANASYLRERITRWIDEE	574
Db	239	RPRTTVLGNLPGTWVYRMDQYAAAQTVPGLVLRVDAPVYFANASYLRERISRWIDEE	298
Qy	575	ERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQLVLVNPVSEVMKKLN	634
Db	299	ERTKSQCEMGVRYVWLDMGAIGSIDTSGTSMLELNKSLDRRGMQIVLANPGSEIMKKLD	358
Qy	635	KSKFQNLGKKWLYLTYVEEAVGACNFNLASKTNPKXD	672
Db	359	SSKVLQIGHWVFPPTVGEAVASCDYVLHSHKPGMAXD	396

Search completed: August 31, 2005, 11:53:40  
Job time : 59 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2005, 08:46:03 ; Search time 35 Seconds  
(without alignments)  
1869.354 Million cell updates/sec

Title: US-10-762-049-18  
Perfect score: 3475  
Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGWNV 680

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2642.5	76.0	658	2 T48901	sulfate transport
2	2641.5	76.0	658	2 T49069	sulfate transport
3	2638.5	75.9	703	2 T51161	hypothetical prote
4	2575.5	74.1	646	2 T48902	sulfate transport
5	2314	66.6	646	2 T01079	sulfate transport
6	1898.5	54.6	631	2 B86365	probable sulphate
7	1796	51.7	683	2 B86354	protein F282.22 [i
8	1758.5	50.6	719	2 D96809	protein F28K19.22
9	1752	50.4	649	2 T00946	probable sulfate t
10	1751.5	50.4	649	2 T51839	sulfate transport
11	1749.5	50.3	660	2 T04416	sulfate transport
12	1735.5	49.9	662	2 S51764	sulfate transport
13	1722.5	49.6	667	2 S51763	sulfate transport
14	1507	43.4	677	2 T50022	sulfate transport
15	1503	43.3	644	2 S51765	sulfate transport
16	1392.5	40.1	658	2 S74246	sulfate transport
17	998	28.7	485	2 S34800	sulfate transport
18	840.5	24.2	750	2 B90137	sulfate permease [
19	690.5	19.9	573	2 D83440	probable sulfate t
20	687.5	19.8	233	2 T01205	sulfate transport
21	658	18.9	566	2 T47633	high affinity sulf
22	658	18.9	605	2 G70409	high affinity sulf
23	658	18.9	840	2 T39116	probable sulfate p
24	649.5	18.7	877	2 T40413	sulfate permease
25	637.5	18.3	739	2 A54808	diastrophic dyspla
26	618.5	17.8	553	2 AE1140	probable sulfate t
27	615	17.7	560	2 F70688	hypothetical prote
28	613.5	17.7	553	2 A11498	probable sulfate t
29	598	17.2	749	2 T23628	hypothetical prote

30	586.5	16.9	764	2 A47456	down-regulated in
31	578	16.6	893	2 S64926	probable membrane
32	571.5	16.4	703	2 A49994	sulfate transport
33	565.5	16.3	611	2 T27820	hypothetical prote
34	563	16.2	754	2 S52816	probable membrane
35	562	16.2	592	2 B82498	sulfate permease f
36	559.5	16.1	859	2 S46176	sulfate transport
37	554.5	16.0	700	2 T23629	hypothetical prote
38	510	14.7	650	2 T16077	hypothetical prote
39	481.5	13.9	667	2 T41306	probable sulfate t
40	467	13.4	544	2 AH1544	transport protein
41	463.5	13.3	758	2 T25751	hypothetical prote
42	451	13.0	537	2 F97546	hypothetical prote
43	451	13.0	537	2 AB2766	hypothetical prote
44	448	12.9	517	2 C83642	sulfate permease [
45	430	12.4	541	2 A11186	probable sulfate t
					transport proteins

ALIGNMENTS

RESULT 1

T48901 sulfate transporter ATST1 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004

C;Accession: T48901

R;Sohlberg, L.E.; Sussex, I.M.

A;Title: Nucleotide sequence of a cDNA (Accession No. U52970) encoding a Cys proteinase f

A;Reference number: Z15626

A;Accession: T48901

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-658 <SOH>

A;Cross-references: UNIPROT:Q9SVL3; EMBL:D89631; PIDN:BAA21657.1

A;Experimental source: Landsberg erecta; tissue type above-ground parts

C;Genetics:

A;Gene: ATST1

C;Superfamily: sulfate transport protein

Query Match 76.0%; Score 2642.5; DB 2; Length 658;  
Best Local Similarity 76.0%; Pred. No. 2e-174;  
Matches 501; Conservative 80; Mismatches 75; Indels 3; Gaps 3;

Qy 24 MGSVDYEYPLGMNFERVHQ-VEVPPQPFKSLKYSKXETFPDDPLRQFNKPASKKF 82

Db 1 MGTEDYTFQGAELHRRHHTVEAPQPFLKSLQYSVKETLFPDDPFRQFNQNASKRF 60

Qy 83 MLGLQFFPIFEWAPKYTFQFLKADLIAGITIASLAIPOGISYAKLANLPPILGLYSSFI 142

Db 61 VLGLKYFLPIFEWAPRYNLKFFKSDLIAGITIASLAIPOGISYAKLANLPPILGLYSSFV 120

Qy 143 PPLIYAMGSSRDLAGVTAVGSLMGSLSNADVDPKLYLHLAFTATTLFAGVFOAA 202

Db 121 PPLIYAVLGSSRDLAGVTAVASLLTGAMLSKEVDAEKOPKLYLHLAFTATTFAGVLEAS 180

Qy 203 LGFLRLGLIVDFLSHATIIIGFMGGAATVCLQKLSILGLEHFTHGADIIISVMRSVPTQT 262

Db 181 LGIFRLGFIIVDFLSHATIVGFMGAATVVSLOQLKGI FGLKHFTDSTDVIVMRSVFSQT 240

Qy 263 HEWRWSAVLGCVPFIFFLLSTRYFSKRRPFPFWSAMAPLTSVLGSLLYFYTHAEKHGV 322

Db 241 HEWRWSGVLGCGFLPFLSTRYFSIKKPKFFWVAAMAPLTSVLGSLLYFYTHAEKHGV 300

Qy 323 EVIGELKGLNPPSLTNLVPSYMTTAVKTVGVGIIISLAEGIAVGRSFAMKKNYNIDG 382

Db 301 QVIGDUKKGLNPLUGSDLIFTSPYMTAVKTVGIIISLAEGIAVGRSFAMKKNYNIDG 360

Qy 383 NKEMIAIGTMNVVGSFSTCYLTGTPFSRAVNNAGCKTAASNIIMSLAVMLTLLFTPL 442

Db 361 NKEMIAFGMNVVGSFSTCYLTGTPFSRAVNNAGCKTAMSNIVMAIVMFTLLFTPL 420

QY	443	FHYTPLVLSAIIIVSAMGLIDYEAAILHLFKVDKDFVVCMSAYIGVFGSVVEIGLVIAI	502
Db	421	FHYTPLVLSAIIISAMGLIDYQAAIHLWKVDKDFFLVCMSAYGVFGSVVEIGLVAV	480
QY	503	VISVLRLVLLFIARPRFTVLGNIPNSVIYRNVEHYQNAKHVPQMLILEIDAPIYFANASYL	562
Db	481	AIISARLLLFVSRPKTAVKGNIPNSMIYRNTQYPSRTPVFGILLILEIDAPIYFANASYL	540
QY	563	RERITRWIDEERIKATGETSLQYVIIDMSAVGNIDTSGISMLBEVKKITERRELQVL	622
Db	541	RERIIRWIDEERVKQSGESSLQYIILDMASVGNIDTSGISMMVEIKKVIDRRALKVL	600
QY	623	VNPVSEVMKLNKSKF-QNHGLGKKWYLTVEEAVGACNFNLASKTNPKKDETEGNNV	680
Db	601	SNPKGEVVKLTRSKTIGDHLGKEMFLTVGEAVEACSYMLHTFKTEP-ASKNEPNNV	658
RESULT 2			
T49069			
sulfate transporter (AT5T1) - Arabidopsis thaliana			
N/Alternate names: protein F4F15.10			
C/Species: Arabidopsis thaliana (mouse-ear cress)			
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004			
C/Accession: T49069			
R/Alcaraz, J.P.; Clabault, G.; Cottet, A.; Mache, R.; Mewes, H.W.; Rudd, S.; Lemcke, K.; submitted to the Protein Sequence Database, April 2000			
A/Reference number: Z25015			
A/Status: preliminary			
A/Molecule type: DNA			
A/Residues: 1-658 <AC>			
A/Cross-references: UNIPROT:Q9SV13; EMBL:AL049711; GSPDB:GN00061; ATSP:F4F15.10			
A/Experimental source: cultivar Columbia; BAC clone F4F15			
C/Genetics:			
A/Gene: ATSP.F4F15.10			
A/Map position: 3			
A/Introns: 116/1; 185/2; 242/3; 264/3; 302/3; 341/3; 384/1; 479/3; 543/2; 572/1; 597/3			
C/Superfamily: sulfate transport protein			
Query Match 76.0%; Score 2641.5; DB 2; Length 658;			
Best Local Similarity 75.9%; Pred. No. 2.4e-174;			
Matches 500; Conservative 81; Mismatches 75; Indels 3; Gaps 3;			
QY	24	MGSVDYEYPLGMNFRVHQ-VEVPPPPFFKSLKYLKTEFPDPPDLQFKKNKASKKF	82
Db	1	MGTEDTFFQGAELHRRHHTVEAPQPQPKLSQYSVKETLFPDPPFQFKQNQASRKF	60
QY	83	MLGLQFPFFIFEWAPKTYTQFLKADLIAGITIASLAIPOGISYAKLANLPPILGLYSFI	142
Db	61	VLGLKYFLPIFEWAPRYNLKPKFSDLIAGITIASLAIPOGISYAKLANLPPILGLYSFV	120
QY	143	PPLIYAMGSSRDLAGVTAVGSLMGSLSNVDPNEDPKLYLHLAFTATLFAGVFOAA	202
Db	121	PPUVYAVLGSSRDLAGVTAVASLLTGAMLSKEVDAEKDPKLYLHLAFTATLFAGVLEAS	180
QY	203	LGLFRLGLIVDFLSHATIIIGFMGGAATVVCLOQLKSLILGLEHFTHGADIISVMRSVFTQT	262
Db	181	LGIFRLGFIIVDFLSHATIVGFMGGAATVVSLOQLKGI FGLKHFTDSTDVISMRSVFSQT	240
QY	263	HEWRWESAVLGCVFIFLLSTRYFSKRPFRFFWVSAMAPLTSVILGSLLYVFFTHAEKHGV	322
Db	241	HEWRWESVGLCGFLFLLSTRYFSIKPKPFFWVAAMAPLTSVILGSLLYVFFTHAERHGV	300
QY	323	EVIGELKKGLNPNLSNLNVPVSYMTAVKTVGVGIISLAEGIAVGRSFAMKYNINIDG	382
Db	301	QVIGDLKKGLNPLSGSDLIFTSPYMSAVKTVGTLTGIIAALAEVAVGRSFAMKYNINIDG	360
QY	383	NKEMIAIGTMNVVGSFTSCYLLTTPGFSRSVAVNAGCKTAASNIIMSLAVMLTLFLTEL	442
Db	361	NKEMIAFGMWNIVGSFTSCYLLTTPGFSRSVAVNAGCKTAMSNIVMAIWMFLTLFLTEL	420
QY	443	FHYTPLVLSAIIIVSAMGLIDYEAAILHLFKVDKDFVVCMSAYIGVFGSVVEIGLVIAI	502

Db	421	FHYTPLVLSAIIISAMGLIDYQAAIHLWKVDKDFFLVCMSAYYGVFGSVVEIGLVAV	480
QY	503	VISVLRLVLLFIARPRFTVLGNIPNSVIYRNVEHYQNAKHVPQMLILEIDAPIYFANASYL	562
Db	481	AIISARLLLFVSRPKTAVKGNIPNSMIYRNTQYPSRTPVFGILLILEIDAPIYFANASYL	540
QY	563	RERITRWIDEERIKATGETSLQYVIIDMSAVGNIDTSGISMLBEVKKITERRELQVL	622
Db	541	RERIIRWIDEERVKQSGESSLQYIILDMASVGNIDTSGISMMVEIKKVIDRRALKVL	600
QY	623	VNPVSEVMKLNKSKF-QNHGLGKKWYLTVEEAVGACNFNLASKTNPKKDETEGNNV	680
Db	601	SNPKGEVVKLTRSKTIGDHLGKEMFLTVGEAVEACSYMLHTFKTEP-ASKNEPNNV	658
RESULT 3			
T51161			
hypothetical protein [imported] - Arabidopsis thaliana			
C/Species: Arabidopsis thaliana (mouse-ear cress)			
C/Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 09-Jul-2004			
C/Accession: T51161			
R/Comella, P.; Wu, H.J.; Laudie, M.; Berger, C.; Cooke, R.; Delseny, M.; Grellet, F.; Plant Mol. Biol. 41, 687-700, 1999			
A/Title: Fine sequence analysis of 60 kb around the Arabidopsis thaliana AtEm1 locus on			
A/Reference number: Z24835; MUID:20108326; PMID:10645728			
A/Accession: T51161			
A/Status: preliminary; translated from GB/EMBL/DBJ			
A/Molecule type: DNA			
A/Residues: 1-703 <COM>			
A/Cross-references: UNIPROT:Q9SV13; EMBL:AF049236; PIDN:AAC14417.1			
C/Genetics:			
A/Map position: 3			
A/Introns: 161/1; 230/2; 287/3; 309/3; 347/3; 386/3; 429/1; 524/3; 588/2; 617/1; 642/3			
C/Superfamily: sulfate transport protein			
Query Match 75.9%; Score 2638.5; DB 2; Length 703;			
Best Local Similarity 75.7%; Pred. No. 4.1e-174;			
Matches 499; Conservative 82; Mismatches 75; Indels 3; Gaps 3;			
QY	24	MGSVDYEYPLGMNFRVHQ-VEVPPPPFFKSLKYLKTEFPDPPDLQFKKNKASKKF	82
Db	46	MGTEDTFFQGAELHRRHHTVEAPQPQPKLSQYSVKETLFPDPPFQFKQNQASRKF	105
QY	83	MLGLQFPFFIFEWAPKTYTQFLKADLIAGITIASLAIPOGISYAKLANLPPILGLYSFI	142
Db	106	VLGLKYFLPIFEWAPRYNLKPKFSDLIAGITIASLAIPOGISYAKLANLPPILGLYSFV	165
QY	143	PPLIYAMGSSRDLAGVTAVGSLMGSLSNVDPNEDPKLYLHLAFTATLFAGVFOAA	202
Db	166	PPUVYAVLGSSRDLAGVTAVASLLTGAMLSKEVDAEKDPKLYLHLAFTATLFAGVLEAS	225
QY	203	LGLFRLGLIVDFLSHATIIIGFMGGAATVVCLOQLKSLILGLEHFTHGADIISVMRSVFTQT	262
Db	226	LGIFRLGFIIVDFLSHATIVGFMGGAATVVSLOQLKGI FGLKHFTDSTDVISMRSVFSQT	285
QY	263	HEWRWESAVLGCVFIFLLSTRYFSKRPFRFFWVSAMAPLTSVILGSLLYVFFTHAEKHGV	322
Db	286	HEWRWESVGLCGFLFLLSTRYFSIKPKPFFWVAAMAPLTSVILGSLLYVFFTHAERHGV	345
QY	323	EVIGELKKGLNPNLSNLNVPVSYMTAVKTVGVGIISLAEGIAVGRSFAMKYNINIDG	382
Db	346	QVIGDLKKGLNPLSGSDLIFTSPYMSAVKTVGTLTGIIAALAEVAVGRSFAMKYNINIDG	405
QY	383	NKEMIAIGTMNVVGSFTSCYLLTTPGFSRSVAVNAGCKTAASNIIMSLAVMLTLFLTEL	442
Db	406	NKEMIAFGMWNIVGSFTSCYLLTTPGFSRSVAVNAGCKTAMSNIVMAIWMFLTLFLTEL	465
QY	443	FHYTPLVLSAIIIVSAMGLIDYEAAILHLFKVDKDFVVCMSAYIGVFGSVVEIGLVIAI	502
Db	466	FHYTPLVLSAIIISAMGLIDYQAAIHLWKVDKDFFLVCMSAYYGVFGSVVEIGLVAV	525
QY	503	VISVLRLVLLFIARPRFTVLGNIPNSVIYRNVEHYQNAKHVPQMLILEIDAPIYFANASYL	562

Db 526 AISIARLLLVSRPKTAVKGNIPNSMIYRTEQYPPSSRTVPGILLIILEIDAPIYFANASYL 585  
Qy 563 REIRITWIDEEERIKATGETSLOQYIILDMASVGNIDTSGISMLEEVKKITERRELQVLV 622  
Db 586 REIRITWIDEEERVKQSGESSLOQYIILDMASVGNIDTSGISMMVEIKKVIDRRAKLVL 645  
Qy 623 VNPVSEVMKLNKSKF-QNHGLGKKWIYLTVEEAVGACNFNLASKTNPKKDETEGNV 680  
Db 646 SNPKGEVVKLTRSKFIGDHLGKEMFLTVGEAVEACSYMLHTFKTEP-ASKNEPNV 703

RESULT 4  
T48902  
sulfate transporter AST12 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
C:Accession: T48902  
R:Takahashi, H.; Sasaki, N.; Kimura, A.; Watanabe, A.; Saito, K.  
Plant Physiol. 121, 686, 1999  
A:Title: Identification of two leaf-specific sulfate transporters in Arabidopsis thaliana  
A:Reference number: Z25001  
A:Accession: T48902  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-646 <KAP>  
A:Cross-references: UNIPROT:Q9SV13; EMBL:AB012048; PIDN:BAAZ5175.1  
A:Experimental source: ecotype Columbia  
C:Genetics:  
A:Map position: 3  
A:Introns: 116/1; 185/2; 242/3; 264/3; 302/3; 329/3; 372/1; 467/3; 531/2; 560/1; 585/3  
C:Superfamily: sulfate transport protein

Query Match 74.1%; Score 2575.5; DB 2; Length 646;  
Best Local Similarity 74.4%; Pred. No. 8.2e-170;  
Matches 490; Conservative 80; Mismatches 74; Indels 15; Gaps 4;

Qy 24 MGSVDYEPYLCMNFERVHQ-VEVPPPPQPFKSLKYSLEKTFPPDDPLROFKKPKASKF 82  
Db 1 MGTEDTYPPQGAELHRRHHTVPAQPQPFKSLQYSKTELPDPDPFPFQKNQNASRK 60  
Qy 83 MLGLQPFPIFEWAPKYTFQFLKADLIAGITIASLAIPQGISYAKLANLPPIGLYSSFI 142  
Db 61 VLGLKYFLPIFEWAPRYNLKFKSDLIAGITIASLAIPQGISYAKLANLPPIGLYSSFV 120  
Qy 143 PPIIYAMGSSRDLAGVTGAVGSLMGSMLSNAVDNEDPKLYLHLAFTATLPAFVQAA 202  
Db 121 PPLVYAVLGSSRDLAGVTGAVASLLTGAMLSKEVDAEKDPKLYLHLAFTATFPAGVLEAS 180  
Qy 203 LGLFRGLIYDFLSHATIIQFMGGAATVCLQOLKSLGLEHETHGADIISVMRSVFTQT 262  
Db 181 LGIHFRGLIYDFLSHATIIQFMGGAATVWSLQOLKGIHGLHFTDSDIVSMRSVFSQT 240  
Qy 263 HEWRWESAVLGCVFIFFLLSTRYFSKKRPFFWVSAMAPLTSVILGSLVYFTHAEKHGV 322  
Db 241 HEWRWESVGLCGFLPFLSTRIFSKKKPFVWVAMAPLTSVILGSLVYFTHAEKHGV 300  
Qy 323 EVIGELKKGSLNPLSLNLVFSVPMYMTAVKTGIVGVIISLAEGIAVGRSPAMYKNYIDG 382  
Db 301 QV-----GSDLIFTSPVMTAVKTGLITGIILAEGVAVGRSPAMFKNYIDG 348  
Qy 383 NKEMIALGTWNVGSGFTSCYLTGPPSRSAVNNAGCKTAASNIIMSLAWMLTLFLTP 442  
Db 349 NKEMIAFGMNNVGSFTSCYLTGPPSRSAVNNAGCKTAMSNVMAIYVMTLFLTP 408  
Qy 443 FHTPLVLSAIIYSAMGLIDYEAAIHLFKVDKDFVWCMSAYIGVFGSVIGLVI 502  
Db 409 FHTPLVLSAIIYSAMGLIDYQAAIHLWKVDKDFVWCMSAYGVVFGSVIGLVAV 468  
Qy 503 VISLVRLLFIARPTFVLGNIPINSVIYRNVHQAQNAKHPGMLIILEIDAPIYFANASYL 562  
Db 469 AISIARLLLVSRPKTAVKGNIPNSMIYRTEQYPPSSRTVPGILLIILEIDAPIYFANASYL 528  
Qy 563 REIRITWIDEEERIKATGETSLOQYIILDMASVGNIDTSGISMLEEVKKITERRELQVLV 622

Db 529 REIRITWIDEEERVKQSGESSLOQYIILDMASVGNIDTSGISMMVEIKKVIDRRAKLVL 588  
Qy 623 VNPVSEVMKLNKSKF-QNHGLGKKWIYLTVEEAVGACNFNLASKTNPKKDETEGNV 680  
Db 589 SNPKGEVVKLTRSKFIGDHLGKEMFLTVGEAVEACSYMLHTFKTEP-ASKNEPNV 646

RESULT 5  
T01079  
sulfate transport protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T01079; T52295  
R:Kaplan, N.; Johnson, D.; Schutz, K.; Gnoj, L.; Hoffman, J.; Till, S.; de la Bastide, M.  
hi, M.; Martienssen, R.; Chen, E.Y.; Wilson, R.; McCombie, W.R.  
Submitted to the EMBL Data Library, November 1998  
A:Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.  
A:Reference number: Z14248  
A:Accession: T01079  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-646 <KAP>  
A:Cross-references: UNIPROT:O04289; EMBL:AC002330; NID:g2262135; PID:g2262137  
A:Experimental source: cultivar Columbia  
R:Takahashi, H.; Sasaki, N.; Kimura, A.; Watanabe, A.; Saito, K.  
Plant Physiol. 121, 686, 1999  
A:Title: Identification of two leaf-specific sulfate transporters in Arabidopsis thaliana  
A:Reference number: Z25001  
A:Accession: T52295  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-646 <TAK>  
A:Cross-references: EMBL:AB004060; PIDN:BAAZ20282.1  
A:Experimental source: strain Columbia  
C:Genetics:  
A:Map position: 4  
A:Introns: 107/1; 176/2; 233/3; 255/3; 293/3; 332/3; 375/1; 470/3; 534/2; 563/1; 588/3  
A:Note: T10P11.3  
C:Superfamily: sulfate transport protein  
C:Keywords: sulfate transport; transmembrane protein

Query Match 66.6%; Score 2314; DB 2; Length 646;  
Best Local Similarity 67.2%; Pred. No. 8.7e-152;  
Matches 430; Conservative 104; Mismatches 102; Indels 4; Gaps 2;

Qy 42 HOVEVPPPPQPFKSLKYSLEKTFPPDDPLRQFNK-PASKKFMGLQFPFPIFEWAPKY 100  
Db 10 HQVEIPPPQPFKSLKNTLNEILFADDPFRIRNESKTSKKIELGLRHVPPILEWARGYS 69  
Qy 101 FQFLKADLIAGITIASLAIPQGISYAKLANLPPIGLYSSFIPLIYAMGSSRDLAGVT 160  
Db 70 LEYKSDVISGITIASLAIPQGISYAQLANLPPIGLYSSLPVPLVYAIMGSSRDLAGVT 129  
Qy 161 VAVGSLMGSMLSNAVDNEDPKLYLHLAFTATLPAFVQAAALGLFRGLIYDFLSHATI 220  
Db 130 VAVASLLTAAMLKEVNAVNVNPKLYLHLAFTATFFAGLMQTCIGLGLRLGFWVEILSHAAI 189  
Qy 221 IQFMGGAATVCLQOLKSLGLEHETHGADIISVMRSVFTQTHWRWESAVLGCVIFFL 280  
Db 190 VQFMGGAATVCLQOLKGLGLHHTHSTDIVTVLRSIFSQSHWRWESVGLGCCFLIFL 249  
Qy 281 LSTRYFSKKRPFFWVSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGSLNPLSLNL 340  
Db 250 LTTKYSKKRPKLFWISAMSPVSVIFGTIFLYFLHQFHGIQFIGELKKGINPPSITHL 309  
Qy 341 VVVSPPMTAVKTGIVGVIISLAEGIAVGRSPAMYKNYIDGNKEMIAIGTWNVGSFTS 400  
Db 310 VFTPPYVMLAKVGIITGVIALAEGIAVGRSPAMYKNYIDGNKEMIAFGMNNILGSFSS 369  
Qy 401 CYLTTCPPSRSAVNNAGCKTAASNIIMSLAWMLTLFLTPHYTPLVLSAIIYSAML 460  
Db 370 CYLTTCPPSRSAVNNAGCKTALSNNVMAVAVATLFLTPFFYTPLVVLSIIAAML 429

Qy 461 GLIDYEAAIHLFKVDKEDFVVCMSAYIGVFGSGVEIGLVIAIVISLVRVLLFIARPRTFV 520  
Db 430 GLVDYEAAIHLWKDKEDFFVCLISAYIGVFGTIEIGLISVGSVNRVLVLFVGRPKIYV 489  
Qy 521 LGNIPNSVIVRNVEHYONAKHPGMLILEIDAPTYFANASYLRIERTWRIDEEERIKAT 580  
Db 490 MGNIONSEIYRNIBHYPOATIRSSLLTLHDGPIYFANSTYDRIDRIGRWIDDEEDKLRTS 549  
Qy 581 GETSLOQVVIIDMSAVGNIDTSGISMLSEVKKITERRRELQGLVLPNPVSEVMKLNKSKFQN 640  
Db 550 GDLSLOQVVIIDMSAVGNIDTSGISMLSEELNKLGRRELKLVIANPGEAVWKKLSKSTFIE 609  
Qy 641 HLGKKWLYLTVEEAVGACNFNLRSKNTNPKKDETEGWNV 680  
Db 610 SIGKERIYLTVAEEAAACDFMLHTAKPDSVPVE---FNNV 646

RESULT 6

B86365  
Probable sulphate transporter protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: B86365  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chan, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizart, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: B86365  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-631 <STO>  
A:Cross-references: UNIPROT:Q98XS2; GB:A5005172; NID:g2829902; PIDN:AAC00610.1; GSPDB:GN  
C:Genetics:  
A:Map position: 1  
C:Superfamily: sulfate transport protein

Query Match 54.6%; Score 1898.5; DB 2; Length 631;  
Best Local Similarity 57.1%; Pred. No. 3.7e123;  
Matches 353; Conservative 125; Mismatches 139; Indels 1; Gaps 1;

Qy 41 VHQVEVPPQPFKSKLSYKETFDDPLQPKNKPKASKKFMGLQFFPPFEPWAPKYT 100  
Db 3 VHKVAPPHKSTVAKLTKLKETFFDDPLQPKNRKTLIRAAQYIFPILQWCPEYS 62  
Qy 101 FQELKADLIAGITIASLAIPOGISYAKLANLPILGLYSSFIPPLIYAMGSSRDLAGVT 160  
Db 63 FSLKSDVSVGLTIASLAIPOGISYAKLANLPILVGLYSSFVPLVAVYLVGSSRDLAGVP 122  
Qy 161 VAVGSLMLGMLSNVDPNDPDKLYLHLAFTATLFAGVFOAALGLFRGLIVDFLSHATI 220  
Db 123 VSTASLILGMLRQVSPVDDVFLQLAPSSFFAGLFOASGLIIRLGFIDFLSKATL 182  
Qy 221 IGFMGGAATVVCLOQLSKILGLEBHTHGADIISVMRSVFTQTHERWESAVLGCVFIFFL 280  
Db 183 IGFMGGAATVVCLOQLSKILGLEBHTHGADIISVMRSVFTQTHERWESAVLGCVFIFFL 242  
Qy 281 LSTRYFSKKEPRFVWSAMAPLTSVILGSLLVYFTHAEKHGVEVIGELKKLNPPSLTNL 340  
Db 243 LSTRHLISMKKPKLFWWSAGAPLSVIVSTLLVFVFAERHGISVIGKLEGLNPPSWNNML 302  
Qy 341 VFVSPYMTTAVKTVGVVGIISLAEGIAVGRSFAMYNKYNIDGNKEMIAICTMNVVGSFTS 400  
Db 303 QFHGSHLALVAKTGLVTGIVSLTEGIAVGRTPAALKKNYHVDGNKEMIAIGLMNVGSATS 362  
Qy 401 CYLTGTPFSRSVNNYAGCKTAASNTIMSLVAMVMTLLFLTPFLPHYTPLVVLISAIIVSAML 460

[illegible]

Db 326 TFFVYITRADKKGVQIVKHLDKGLNPSLSRLIYFSGDYLLKGFRIQVVGSHVALTEAVAI 385  
Qy 369 GRSFAMYNKYNIDGNKEMIAIGTNVVGSTSCYLTTGPFSSRSVAVNAGCKTAASNIIM 428  
Db 386 GRTFAAMKDYQIDGNKEMALGAMNVIGSMSTCVSTGTSFSSRSVAVNFMAGCQTAVSNIIM 445  
Qy 429 SLAVMLTLLELTPPLPHYTTPVLVLSAIIIVSAMGLIDYEAHILFKVDKFPVVCMSAYIG 488  
Db 446 SIVVLLTLELTPPLFKYTPNAILAAIIINAVIPLDVNVATILPKIDKLDFAVACMGAFPG 505  
Qy 489 VVFGSVEIGLIIAGVSGAKILLQVTPRTAILGKIPGTSVYRNINQYPEATRIPGVLT 548  
Db 506 VIFVSVEIGLIIAGVSGAKILLQVTPRTAILGKIPGTSVYRNINQYPEATRIPGVLT 565  
Qy 549 EIDAPIYFANASYLRERITWIDDEERIIRKATGETSLQYVIIDMSAVGNIDTSGISMLEE 608  
Db 566 RVDSAIYFSNYSVYRRIQRLWIDDEEMVZEARLPRIQFLIEMSPVTDIDTSGIHALED 625  
Qy 609 VKKITERRELQVLVNPVSEVMKLNKSKFQNHGKKWIIYLTVEEAVGACN 659  
Db 626 LYKSLQKRDIOQLVLANPGPPVINKLVSHFADLGHDKIFLTVAEAVDSCS 676  
RESULT 8  
D96809  
Protein F28K19.22 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: D96809  
R;Theologas, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,  
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: D96809  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-719 <STO>  
A;Cross-references: UNIPROT:Q9MAX3; GB:AE005173; NID:96573765; PIDN:AAF17685.1; GSPDB:GN  
C;Genetics:  
A;Gene: F28K19.22  
A;Map position: 1  
C;Superfamily: sulfate transport protein  
Query Match 50.6%; Score 1758.5; DB 2; Length 719;  
Best Local Similarity 49.9%; Pred. No. 2e-113;  
Matches 342; Conservative 116; Mismatches 157; Indels 71; Gaps 5;  
Qy 42 HQVEVPPQPFKSLKYSLETFFDDPLQFNKNPKASKFMLGLQFPFIFWAKYTF 101  
Db 28 HKVGIPEKQNFDMYTFETFFDDPLRDFKQPKSKQFMLGLQSFVFPVDFWGRNYTF 87  
Qy 102 QFLKADLIAGITIASLAIPQ-----GISVAKLANLPPIILGLYSSP 141  
Db 88 KFRGDLISGLTASLCIPQVKNLSSTCTSYLLVLYVDIGVAKLANDPKYGLYSSP 147  
Qy 142 IPELLYAMGSSRLDAGTVAVGSLMLGMSLNAVDNEDPKLYLHLAFTATLPAQVFOA 201  
Db 148 VPPLVACMGSSRDIAIGPVAVVSLLGTLRLAIEDNTSPDEYLRLAFTATFAGITEA 207  
Qy 202 ALGLFRGLIVDFLSHATIIIGFMGGAATVVCLOOLKSIILGLEHFTHGADIISVWRSVFTQ 261  
Db 208 ALGFRFRGLIDFLSHAHVAVGFMGGAATIALQOLKGLGFKTKTKTDIISVLESVFK 267  
Qy 262 THE-WRWSAVLGCVTFEFLLSLTRYF-----SKRPP 291  
Db 268 AHGWNWVILIGASFLTLTSLTIIVRHISINKTSKFIILCLDLFLTSLDMLKQKCKSK 327

Qy 292 RFPWVGAMAPLTSVILGSLLYVTFHAEKHGVE-----VIGELKKGL 332  
Db 328 KLFVWPAIAPLISVIVSTFTFYITRADKQGVQIVRSQPLTSLFRFKQFVVVVKHLDOGI 387  
Qy 333 NPESLTNLVFPSPYMTAVKTGIVVGIISLAEGIAVCRSPAMKYNIDGNKEMIAIGTM 392  
Db 388 NPSSFHLIYFTGDNLAIRIGVVGWGMVALT--VTGRTFAAMKDYQIDGNKEMVAGMM 445  
Qy 393 NVVGSTSCYLTTGPFSSRSVAVNAGCKTAASNIIMSLAVMLTLLELTPPLPHYTTPVLVLS 452  
Db 446 NVVGSMSSCVATGTSFSSRSVAVNFMAGCQTAVSNIINSIVVLLTLELTPPKYTPNAILA 505  
Qy 453 AIIIVSAMGLIDYEAHILFKVDKFPVVCMSAYIGVVGSGVEIGLIIAVISVLRVLLF 512  
Db 506 AIIINAVIPLIDITQAAAILFKVDKLDFAICAGFFGVIFVSVIEGLIIAVSISPAKILLQ 565  
Qy 513 IARPRFVLGNTPNSVYRNVEHYQNAKHVPGLMLILEIDAPIYFANASYLRERITWIDE 572  
Db 566 VTRPRTAVLGNIPRTSVYRNIOQYPEATVWPGVLTIRVDSAIYFSNYSVYRRIQRLWHE 625  
Qy 573 EERIIRKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKK 632  
Db 626 EEEKVKAASLPRIQFLIEMSPVTDIDTSGIHALEDLYKSLQKRDIOQLVLANPGPLVIGK 685  
Qy 633 LNKSFKQNHGKKWIIYLTVEEAVGAC 658  
Db 686 LHLSHADPMDLQDNIIYLTVADEAVAC 711  
RESULT 9  
T00946  
Probable sulfate transport protein T3F12.7 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: T00946  
R;Gnoj, L.; Huang, B.N.; Habermann, K.; Hameed, A.; Hasegawa, A.; Jensen, K.; Schutz, K.;  
R.; McCombie, W.R.  
submitted to the EMBL Data Library, October 1997  
A;Description: Arabidopsis thaliana BAC T3F12 from chromosome IV.  
A;Reference number: Z14210  
A;Accession: T00946  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-649 <GNO>  
A;Cross-references: UNIPROT:Q9SAV1; EMBL:AC002983; NID:G2443899; PID:G2565006  
C;Experimental source: cultivar Columbia  
C;Genetics:  
A;Map position: 4  
A;Introns: 102/3; 119/1; 246/3; 306/3; 345/3; 388/1; 483/3; 547/2; 576/1; 601/3  
A;Note: T3F12.7  
C;Superfamily: sulfate transport protein  
C;Keywords: sulfate transport; transmembrane protein  
Query Match 50.4%; Score 1752; DB 2; Length 649;  
Best Local Similarity 52.4%; Pred. No. 4.9e-113;  
Matches 333; Conservative 124; Mismatches 177; Indels 2; Gaps 2;  
Qy 34 GMMNFERVHQ-VREVPPQPFKSLKYSLETFFDDPLQFNKNPKASKFMLGLQFPFI 92  
Db 14 GARNPPVQQRVLAPPKAGLLDKISVVEETFFHDAPLDFKQGTAKKALLQIAQVFI 73  
Qy 93 FEWAPKYTFQFLKADLIAGITIASLAIPQGISYAKLANLPPIILGLYSSSFPIIYAMGMS 152  
Db 74 IGWAREYTLRKFRGDLIAGITIASLCIPQDIGYAKLANDPKYGLYSSFPPIIYAGMGS 133  
Qy 153 SRDLAVGTAVGSLMLGMSLNAVDNEDPKLYLHLAFTATLPAQVFOAALGLFRGLIV 212  
Db 134 SRDIAIGPVAVVSLVLTLCQAVIDPKNPKEDYLRVFTATTFAGIFQAGLGLFRGLI 193  
Qy 213 DFLSHATIIIGFMGGAATVVCLOOLKSIILGLEHFTHGADIISVWRSVFTQT-HEMWRVESAV 271  
Db 194 DFLSHAAVVGFMGGAATIALQOLKGLGFKTKTKTDIVSVWHSVFKNAEHNWQTIIV 253





[illegible]

RESULT 12  
S51764 sulfate transport protein 2, high affinity - Stylosanthes hamata  
C:Species: Stylosanthes hamata  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S51764  
R:Smith, F.W.; Ealing, P.M.; Hawkesford, M.J.; Clarkson, D.T.  
submitted to the EMBL Data Library, October 1994  
A:Description: Cloning and expression of cDNAs that encode sulphate transporters from pl  
A:Reference number: S51763  
A:Accession: S51764  
A:Molecule type: mRNA  
A:Residues: 1-662 <SMI>  
A:Cross-references: UNIPROT:P53392; EMBL:X82256; NID:g607185; PIDN:CAA57711.1; PID:g6071  
A:Experimental source: tissue type root  
C:Genetics:  
A:Gene: SHST2  
C:Superfamily: sulfate transport protein  
C:Keywords: sulfate transport

Query Match	49.9%;	Score	1735.5;	DB 2;	Length	662;			
Best Local Similarity	51.0%;	Pred. No.	6.9e-112;						
Matches	324;	Conservative	131;	Mismatches	167;	Indels	13;	Gaps	3
Qy	22	EDMGSVDYIYPIGMNFERVHQVEVPPQPFKSLKSLKETFPDPDPLRQFNKKPASKK	81						
Db	30	DDTPSLPY-----MHKVGAPPQTTLFQIKHSFNETFPDPKPFNGKQDSGRK	78						
Qy	82	FMLGLQFPFPIFEWAPKPTFQPLXADLTAGTITIASLAIPQISVAKLANLPILGLYVSF	141						
Db	79	FVUGLQYIPIIDWGRHYDLKFKRGDFTAGTITIASLCIPQOLAYAKLANDPWGYLSF	138						
Qy	142	IPPLIYAMWGSSRDLAGVTVAVGSLLMGSMSLNADVNEOPKLYLHLAFTATLFAGVFOA	201						
Db	139	VAPLVYAFMGTSRDTAIGPVAVVSLGLTLLSNEIS-NTKSHDYLRLAFTATTFAGVTOM	197						
Qy	202	ALGLPRLGLIVDFLSHATIIIGPMGGAATVVCLOQLKSLGLHEFTHGADIISVMRSVFQO	261						
Db	198	LLGVCRLGLFLDFLSHAIIIGPMGGAATTIGLQQLKGLGKDFTKNSDIVSVMHVSWSN	257						
Qy	262	THE-WRWESAVIGCVPIPFLLISTRVFSKKRRPFWWSAMAPLTSVLGLSLLVYTFTHAEKH	320						
Db	258	VHKGMMNETIIGLSFLPLATITKYAKNNKLFWSAISPMICVISTEFPVITRADKR	317						

Qy	321	QVEVIGELKGLNPPSLTNLVPSPVMTTAVKTVGLVWGIIISLAEGIAVGRSPFAMVQYNNI	380
Db	318	GVTIVGHKSGVNPSSANEEIFPHGKVLGAGVRGVVAGVLVTEAMAJGRTFPAAMKDYSI	377
Qy	381	DGNKEMIAIGTMNVVGSPTSCLYLTGPPSRSAVNNAGCKTAASNIIMSLAVMLTLLFLT	440
Db	378	DGNKEMVAMGTMNVGSLTSCYVTTGSPSRSAVNNAGCKTAVSNIIMVAIVLLTLLVIT	437
Qy	441	PLFHYTPLVVLASIIIVSAMLGLIIDYEAHILFKVDKDFPVVCMASAYIGVVFGSVEIGLVI	500
Db	438	PLFKYTNNAVLASIIIAAVNVLNIEAMVLLWKDKDFVACMGAPFGVIFKSVIEIGLLI	497
Qy	501	AIIVSLVRLLFIAPRFTVLGNIPNSVIYRNVEHYQNAKHVPGMLLILEIDAPIYFANAS	560
Db	498	AVAISFAKILLQVTRPTAVLGLKLPGTGSVYRNIOQYPKAEQIPGMLIIRVDSAIYFSNSN	557
Qy	561	YLRIRTRWIDEEERIKATGETSLOYVIIDMSAVGNTDTSGISMLEVEKKKITEREQL	620
Db	558	YIKERILRLWIDEGARTSELPETQHLIVEMSPVTDIDTSGIHAFEEELYKTQKREVOL	617
Qy	621	VLNVPVSEVMKKLNKSPQNHGLKKWIVLTVBEAV	655
Db	618	MLANPGPVWIKLHASNLAEILGBDKIELTVADAV	652

RESULT 13  
SS1763

sulfate transport protein 1, high affinity - Stylosanthes hamata  
S11763  
C/Species: Stylosanthes hamata  
C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C/Accession: S51763

C:Accession: S51763  
 F:Smith, F.W.; Ealing, P.M.; Hawkesford, M.J.; Clarkson, D.T.  
 submitted to the EMBL Data Library, October 1994  
 A:Description: Cloning and expression of cDNAs that encode sulphate transporters from pl  
 A:Reference number: S51763  
 A:Accession: S51763  
 A:Molecule type: mRNA  
 A:Residues: 1-667 <SMf>  
 A:Cross-references: UNIPROT:P53391; EMBL:X82255; NID:g607183; PIDN:CAA57710.1; PID:g6071  
 A:Experimental source: tissue type root  
 C:Genetics:  
 A:Gene: SHST1  
 C:Superfamily: sulfate transport protein  
 C:Keywords: sulfate transport

Query Match	49.6%	Score 1722.5	DB 2	Length 667
Best Local Similarity	51.3%	Pred. No. 5.5e-111		
Matches 327	Conservative	125	Mismatches 170	Indels 15
				Gaps 4

Qy	22	EDMGSV	DYBYPL	GMNNFER	VHQV	VPPQ	PPFKSL	KYLS	KETFP	DDPL	ROFK	KNP	ASKK	81
Db	33	DDTTS	LPY-----	MHKV	GTP	PKQL	FQEI	KHS	FNET	FP	KDP	FGFK	DQSGPRK	81
Qy	82	FMGLQ	OFFPPI	BEWAK	YTPQ	LKAD	L	AGT	IAS	LAI	PQGIS	YAK	LANL	141
Db	82	LELGL	QYIP	PILEW	GRHYD	LKFR	GDF	AGT	L	IAS	L	CIPO	LAYAK	141
Qy	142	IPLIY	AMGSS	RDLAV	GVAV	GSILL	MSLS	NA	VD	PNED	PKL	YLH	LAFT	201
Db	142	VAPLV	YAFMG	TGRD	IAIG	BPA	VVUS	ULLG	L	LSN	ELIS	-NTK	SHDY	201
Qy	202	ALGLF	RLG	LIVD	PLSH	ATI	IGPM	GGA	AT	VVCI	QQLK	SL	IGL	251
Db	201	LLGVC	RLG	LIVD	PLSH	ATI	IGPM	GGA	AT	VVCI	QQLK	SL	IGL	251
Qy	260	TOTHE	-WRWES	AVLGC	VFF	LLSTR	YFS	KR	PR	FFW	SAM	APL	TS	311
Db	261	THVHG	GNMET	LIG	LSFL	ILL	TKY	AKK	KL	FW	YS	AI	SP	321
Qy	319	XHGVE	VIGEL	KKGLN	PPS	UTNL	V	SPY	MTT	AVK	TGI	VVG	IS	371
Db	321	KRGVS	YTKH	KSGVN	PSS	ANE	IFF	FGK	YGL	GAC	VR	VGV	AGL	381

QY 379 NIDCNKEMIAIGTMNVVGSFTSCYLTTPGFSRQAVNAGCKTAASNIIMSLAVMLTLFL 438  
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 440  
QY 439 LTLFPHYTPVLVLSAIIIVSAMGLIDYEAIAHLFKVDKDFVFCVMSAYIGVFGSVEIGL 498  
Db ITLEFKYTPNAVLSAIIIAAVNALVNEAMVLLKIDKDFVACMGAFGVIKFSVEIGL 500  
QY 499 VIAIVSLRVLLFIARPRFVLGNTPNSVIYRVEHYQNAKHVPGMLLILEIDAPYFAN 558  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 560  
QY 559 ASYLRETRTWIDEEERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEKKITERREL 618  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 620  
QY 619 QLVIVNPVSVMKLNKSKFQNHGKKWIVLTVEEAV 655  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 657  
RESULT 14  
T50022  
sulfate transporter - Arabidopsis thaliana  
N;Alternate names: protein T31P16.170  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
C;Accession: T50022  
R;Bevan, M.; Zimmermann, W.; Gruenisen, A.; Wambutt, R.; Kalicki, J.; Wohldmann, P.; S  
submitted to the Protein Sequence Database, May 2000  
A;Reference number: Z25027  
A;Accession: T50022  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-677 <BEV>  
A;Cross-references: UNIPROT:O04722; EMBL:AL356332; GSPDB:GN00063; ATSP:T31P16.170  
A;Experimental source: cultivar Columbia; BAC clone T31P16  
C;Genetics:  
A;Gene: ATSP:T31P16.170  
A;Map position: 5  
A;Introns: 132/3; 149/1; 298/3; 375/3; 418/1; 513/3; 578/2; 609/1; 634/3  
C;Superfamily: sulfate transport protein  
Query Match 43.4%; Score 1507; DB 2; Length 677;  
Best Local Similarity 48.2%; Pred. No. 4e-96;  
Matches 300; Conservative 128; Mismatches 183; Indels 12; Gaps 8;  
QY 48 PPQPFKSLKYSUKETFFPDPLRQFN---KPASKKFMGLQFPFPPIFEWAPKTYTFQL 104  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 115  
QY 105 KADLIAGITIASLAIQGISYAKLANLPILGLYSSFIPLIYAMGSSRDIAVGVAVG 164  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 175  
QY 165 SLIIMGSLNADVPNEDPKLYHLAFTATLFAGVQAAALGLRLGLIVDFLSHATIIGFM 224  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 235  
QY 225 GGAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQ-THRWKESAVLCGVFIFFLST 283  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 295  
QY 284 RYFSKRPREFWVSAMAPLTSVLGSLIYVFTAEKHGVEVIGELKKNLPDLTLNVV 343  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 355  
QY 344 SPYMTTAVKTIGIVVGIISLAEGIAVGRSFAMYNKYNIDGNKEMIAIGTMNVVGSFTSCYL 403  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 415  
QY 404 TTGPFSSRQAVNAGCKTAASNIIMSLAVMLTLFLTPHYTPVLVLSAIIIVSAMGLI 463

Db ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 475  
QY 416 ATGSPSRQAVNAGCKTAASNIIMSLAVMLTLFLTPHYTPVLVLSAIIIVSAMGLI 475  
QY 464 DYEAIAHLFKVDKDFVFCVMSAYIGVFGSVEIGLVAIVSLRVLLFIARPRFVLGN 523  
Db NINEAIIHWKDKDFLADIGAFGVLFAVSEIGLVAIVSFAKILISIRPIETLGR 535  
QY 524 IPNSVIYRVEHYQNAKHVPGMLLILEI-DAPYFANASYLRERITRWID--BEERIKAT 580  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 595  
QY 581 GETSLQYVIIDMSAVGNIDTSGISMLEEKKITERRELQVLVNPVSEVMKLNKSKFQ 640  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 655  
QY 641 HLGKKWIVLTVEEAVGACNPNLR 663  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 676  
RESULT 15  
S51765  
sulfate transporter protein 3, low affinity - Stylosanthes hamata  
C;Species: Stylosanthes hamata  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S51765  
R;Smith, P.W.; Ealing, P.M.; Hawkesford, M.J.; Clarkson, D.T.  
submitted to the EMBL Data Library, November 1994  
A;Description: Cloning and expression of cDNAs that encode sulphate transporters from pl  
A;Reference number: S51765  
A;Accession: S51765  
A;Molecule type: mRNA  
A;Residues: 1-644 <SMI>  
A;Cross-references: UNIPROT:P53393; EMBL:X82454; NID:g607187; PIDN:CAA57831.1; PID:g60718  
C;Genetics:  
A;Gene: SHST3  
C;Superfamily: sulfate transport protein  
C;Keywords: sulfate transport  
Query Match 43.3%; Score 1503; DB 2; Length 644;  
Best Local Similarity 47.0%; Pred. No. 7.1e-96;  
Matches 301; Conservative 124; Mismatches 193; Indels 22; Gaps 6;  
QY 33 LGMNNFERVHQ--VEVPPPPQPFKSLKYSUKETFFPDPLRQFNKPASKKFMGLQF-- 88  
Db 4 LGTEQFSERSQVNLSPNPPLTKKFLGLKD-----NKFFTSSSSKKETRAVSPLA 55  
QY 89 -FPPIFEWAPKTYTFQLKADLIAGITIASLAIQGISYAKLANLPILGLYSSFIPLIY 147  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 115  
QY 148 ANMGSSRDIAVGVAVGSLMGSLNADVPNEDPKLYHLAFTATLFAGVQAAALGLPR 207  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 175  
QY 116 ALMGSSREIAIGPVAVSMGLSSLPVKVDPDAHPNDYRNLVFTVTLFAGIFQAFGVLR 175  
QY 208 LGLIVDFLSHATIIGFMGGAATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQTHE--- 264  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 235  
QY 265 ----NRWESAVLCGVFIFFLSTRYFSKRPREFWVSAMAPLTSVLGSLIYVFTAEKH 320  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 295  
QY 321 GVEVIGELKKNLPDLTLNVVSPYMTTAVKTIGIVVGIISLAEGIAVGRSFAMYNKYN 380  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 355  
QY 381 DGNKEMIAIGTMNVVGSFTSCYLTTPGFSRQAVNAGCKTAASNIIMSLAVMLTLFL 440  
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 415  
QY 441 PLFHYTPVLVLSAIIIVSAMGLIDYEAIAHLFKVDKDFVFCVMSAYIGVFGSVEIGL 500  
| :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: |

Db	416	RLYYTPMAILSILSALPGLIDIGEAYHIWKVDKFDLACLGAFFGVLFVSTIEIGLLI	475
Qy	501	AIIVISVLRLVLLFIARPRTFVLGNIPNSVIYRNVEHYQNAKHVPCMLILEIDA-PIYPANA	559
Db	476	ALSISFAKILLQAIRPGVEVLGRIPPTTEAYCDVAQYPMVTTPGILVIRISSGSLCFANA	535
Qy	560	SYLRERITRWI-DREEERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVEVKKITERREL	618
Db	536	GFVREIRLKWVEDEEQDNIEEAAKGRVQAIIDMTDLTNVDTSGILALEELHKLLSRGV	595
Qy	619	QLVLVNPVSEVMKGLNKSQFQNLGKKWYLYLTVEEAVGAC	658
Db	596	ELAMVNPWEVIHKLKVANFVDKIGKERVEITVAEAVDAC	635

Search completed: August 31, 2005, 11:51:58  
Job time : 37 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2005, 06:33:57 ; Search time 141 Seconds  
(without alignments)

2469.602 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLRNTI.....NLRASKTNPKDETEGNNV 680

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2641.5	76.0	658	1	ST31_ARATH	Q98v13 arabidopsis
2	2596.5	74.7	658	2	Q70GU1	Q70GU1 brassica na
3	2326.5	66.9	637	2	Q8H7X7	Q8H7X7 oryza sativ
4	2314	66.6	646	1	ST32_ARATH	Q04289 arabidopsis
5	2274.5	65.5	650	2	Q70BL1	Q70BL1 brassica ol
6	2057	59.2	659	2	Q7XEF8	Q7XEF8 oryza sativ
7	1898.5	54.6	631	1	ST33_ARATH	Q98x82 arabidopsis
8	1880.5	54.1	656	2	Q9AT47	Q9AT47 lycopersico
9	1875	54.0	653	1	ST34_ARATH	Q91W86 arabidopsis
10	1858	53.5	656	2	Q6Z9Y1	Q6Z9Y1 oryza sativ
11	1850.5	53.3	658	2	Q9AT12	Q9AT12 zea mays (m
12	1841.5	53.0	653	1	ST12_ARATH	Q9max3 arabidopsis
13	1836.5	52.8	629	2	Q7X5T0	Q7X610 oryza sativ
14	1830.5	52.7	656	2	Q6ZXB7	Q6ZXB7 brassica ol
15	1811.5	52.1	666	2	Q942E2	Q942E2 oryza sativ
16	1810.5	52.1	655	2	Q93XL1	Q93XL1 brassica na
17	1809.5	52.1	635	2	Q6ZZ94	Q6ZZ94 brassica ol
18	1807.5	52.0	656	1	ST13_ARATH	Q9fep7 arabidopsis
19	1796.5	51.7	651	2	Q94F02	Q94F02 lycopersico
20	1791.5	51.6	657	2	Q9AT48	Q9AT48 lycopersico
21	1787	51.4	634	1	ST35_ARATH	Q941W6 arabidopsis
22	1777.5	51.2	662	2	Q8H7X3	Q8H7X3 oryza sativ
23	1773.5	51.0	662	2	Q8RUC3	Q8RUC3 oryza sativ
24	1770.5	50.9	648	2	Q93YK6	Q93YK6 brassica na
25	1767.5	50.9	657	2	Q9FPQ4	Q9FPQ4 solanum tub
26	1760.5	50.7	660	2	Q43482	Q43482 hordeum vul
27	1760.5	50.7	660	2	Q84XE9	Q84XE9 hordeum vul
28	1752	50.4	649	1	ST11_ARATH	Q9s4Y1 arabidopsis
29	1749.5	50.3	660	2	Q40008	Q40008 hordeum vul
30	1745.5	50.2	662	2	Q8H0K3	Q8H0K3 triticum ae
31	1738.5	50.0	662	2	Q8H2D8	Q8H2D8 aegilops sp

32	1738.5	50.0	662	2	Q9XGB5	Q9XGB5 aegilops ta
33	1735.5	49.9	662	1	SUT2_STVHA	P53392 stylosanthe
34	1725.5	49.7	655	2	Q8H0J8	Q8H0J8 triticum ur
35	1723.5	49.6	660	2	O04001	O04001 sporobolus
36	1722.5	49.6	667	1	SUT1_STVHA	P53391 stylosanthe
37	1720.5	49.5	662	2	Q8H0K2	Q8H0K2 triticum ae
38	1714.5	49.3	655	2	Q8H0K6	Q8H0K6 triticum ae
39	1710	49.2	655	2	Q8H0K4	Q8H0K4 triticum ae
40	1697	48.8	655	2	Q8H2D7	Q8H2D7 aegilops sp
41	1690.5	48.6	655	2	Q8H0K5	Q8H0K5 triticum ae
42	1680	48.3	652	2	Q8H7X4	Q8H7X4 oryza sativ
43	1666.5	48.0	666	2	Q8H0J9	Q8H0J9 triticum ur
44	1653	47.6	649	2	Q9XGB6	Q9XGB6 aegilops ta
45	1563.5	45.0	659	2	Q8LR58	Q8LR58 oryza sativ

#### ALIGNMENTS

RESULT 1

ID	ST31_ARATH	STANDARD;	PRT;	658 AA.
AC	Q9SV13; O23250; O64435; O65025;			
DT	10-OCT-2003 (Rel. 42; Created)			
DT	10-OCT-2003 (Rel. 42; Last sequence update)			
DT	25-OCT-2004 (Rel. 45; Last annotation update)			
DE	Sulfate transporter 3.1 (AST12) (AtST1).			
GN	Name=SULTR3;1; Synonyms=ST1; OrderedLocusNames=AT3G51895;			
GN	ORFNames=F4F15.10, ATEM1.15;			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Landsberg erecta; TISSUE=Aerial parts;			
RA	Yamaguchi Y., Nakamura T., Harada E., Koizumi N., Sano H.;			
RT	"Isolation and characterization of a cDNA encoding a sulfate transporter from Arabidopsis thaliana.";			
RL	(er) Plant Gene Register PGR97-051.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RA	Takahashi H., Sasakura N., Kimura A., Watanabe A., Saito K.;			
RT	"Identification of two leaf-specific sulfate transporters in Arabidopsis thaliana.";			
RL	(er) Plant Gene Register PGR99-154.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=20108326; PubMed=10645728; DOI=10.1023/A:1006395324818;			
RA	Comella P., Wu H.-J., Lauder M., Berger C., Cooke R., Delseeny M.,			
RA	Grellet P.;			
RT	"Fine sequence analysis of 60 kb around the Arabidopsis thaliana AtEm1 locus on chromosome III.";			
RL	Plant Mol. Biol. 41:687-700 (1999).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=21016720; PubMed=11130713; DOI=10.1038/35048706;			
RA	Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unsel M.,			
RA	Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,			
RA	Delseeny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,			
RA	De Simone V., Choiane N., Artiguenave F., Robert C., Brottier P.,			
RA	Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,			
RA	Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,			
RA	Murmbach E., Drzonek H., Erle H., Jordan N., Bangert S.,			
RA	Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,			
RA	Vezi A., D'Angelo M., Pallavicini A., Toppi S., Simionati B.,			
RA	Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,			
RA	Reichert J., Scharfe M., Schoen O., Barges M., Terol J., Clement J.,			
RA	Navarro P., Collado C., Perez-Perez A., Ottenwelder B., Duchemin D.,			

RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,  
RA de Haan M., Maarsee A.C., Alcaraz J.-P., Cottet A., Cassacuberta E.,  
RA Monfort A., Argirio A., Flores M., Liguori R., Vitale D.,  
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
RA Rooney T., Rizzo M., Walte A., Utterback T., Fujii C.Y., Shea T.P.,  
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,  
RA Pai G., Miltscher J., Sellers P., Gill J.E., Feldbiyuni T.V.,  
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,  
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,  
RA Sasamoto S., Kimura T., Ideawara K., Kawashima K., Kishida Y.,  
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuuchi A., Muraki A.,  
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,  
RA Watanabe A., Yamada M., Yasuda M., Tabata S.,  
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
RT thaliana."  
RL Nature 408:820-822(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;  
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.P.,  
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.P.,  
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,  
RA Khan S., Koeseena E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,  
RT "Empirical analysis of transcriptional activity in the Arabidopsis  
RT genome."  
RL Science 302:842-846(2003).  
CC -!- FUNCTION: H(+)/sulfate cotransporter that may play a role in the  
CC regulation of sulfate assimilation.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -!- TISSUE SPECIFICITY: Expressed only in leaves.  
CC -!- INDUCTION: By nitrogen starvation, but not by sulfate starvation.  
CC -!- SIMILARITY: Belongs to the SLC26A/sulf transporter (TC 2.A.53)  
CC family.  
CC -!- SIMILARITY: Contains 1 STAS domain.  
CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous  
CC gene model prediction.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D89631; BAA21657.1; -;  
CC DR EMBL; AB012048; BAA25175.1; ALT SEQ.  
CC DR EMBL; AF049236; AAC14417.1; ALT\_INT.  
CC DR EMBL; AL049711; CAB43310.1; -;  
CC DR EMBL; AY099873; AAM20724.1; -;  
CC DR PIR; T48901; T48901.  
CC DR PIR; T48902; T48902.  
CC DR PIR; T49069; T49069.  
CC DR PIR; T51161; T51161.  
CC DR GeneFarm; 2359; 189.  
CC DR InterPro; IPR02645; STAS.  
CC DR InterPro; IPR001902; Sulfph\_transpt.  
CC DR Pfam; PF01740; STAS; 1.  
CC DR Pfam; PF00916; Sulfate\_transp; 1.  
CC DR ProDom; PD005210; Antisig\_antgnet; 1.  
CC DR TIGRFAMs; TIGR00815; sulf; 1.  
CC DR PROSITE; PS01130; SLC26A; FALSE\_NEG.  
CC DR PROSITE; P550801; STAS; 1.

KW Multigene family; Sulfate transport; Symport; Transmembrane;  
KW Transport.  
FT DOMAIN 1 85 Cytoplasmic (Potential).  
FT TRANSMEM 86 106 Potential.  
FT DOMAIN 107 108 Extracellular (Potential).  
FT TRANSMEM 109 129 Potential.  
FT DOMAIN 130 133 Cytoplasmic (Potential).  
FT TRANSMEM 134 154 Potential.  
FT DOMAIN 155 163 Extracellular (Potential).  
FT TRANSMEM 164 184 Potential.  
FT DOMAIN 185 185 Cytoplasmic (Potential).  
FT TRANSMEM 186 206 Potential.  
FT DOMAIN 207 245 Extracellular (Potential).  
FT TRANSMEM 246 266 Potential.  
FT DOMAIN 267 271 Cytoplasmic (Potential).  
FT TRANSMEM 272 292 Potential.  
FT DOMAIN 293 332 Extracellular (Potential).  
FT TRANSMEM 333 353 Potential.  
FT DOMAIN 354 363 Cytoplasmic (Potential).  
FT TRANSMEM 364 384 Potential.  
FT DOMAIN 385 398 Extracellular (Potential).  
FT TRANSMEM 399 419 Potential.  
FT DOMAIN 420 425 Cytoplasmic (Potential).  
FT TRANSMEM 426 446 Potential.  
FT DOMAIN 447 464 Extracellular (Potential).  
FT TRANSMEM 465 485 Potential.  
FT DOMAIN 486 558 Cytoplasmic (Potential).  
FT TRANSMEM 513 637 STAS.  
FT DOMAIN 637 658 STAS.  
FT CONFLICT 208 208 V -> L (in Ref. 3).  
FT CONFLICT 344 344 V -> I (in Ref. 1).  
SQ SEQUENCE 658 AA; 72748 MW; ED1860A3467FB577 CRC64;  
Query Match 76.0%; Score 2641.5; DB 1; Length 658;  
Best Local Similarity 75.9%; Pred. No. 2.2e-165; Indels 3; Gaps 3;  
Matches 500; Conservative 81; Mismatches 75;  
QY 24 MGSVDYEPGLGMNFERVHQ-VEVPPQPFPFKSLKYLKETFPPDDPLRFQKPKASKKF 82  
DB 1 MGTEDYTPFGAELHRRHTVEAQPQPFKSLQSVKETFPPDDPPFPQKQNASRKF 60  
QY 83 MGLGLOFPPIFWAPKYTFQFLKADLIAGITITASIAIPOGISYAKLANLPPILGYSSFI 142  
DB 61 VLGLKFLPFIWAPRYNLKFFKSLDIAGITITASIAIPOGISYAKLANLPPILGYSSV 120  
QY 143 PPLIYAMGSSRDLAGVTAVGSLMGSLNAPDNPDKLYLHLAFTATLFAVGFOAA 202  
DB 121 PPLVAVLGSSRDLAGVTAVASLITGMLSEVDAEKDPKLYLHLAFTATFFAGVLEAS 180  
QY 203 LGLFLGLIVDFLSHATIGRMGGAATVVCLOOLKSLILGLEHFTHGADIISVMRSVFTQT 262  
DB 181 LGIFRLGFTVDFLSHATIGRMGGAATVVCLOOLKSLILGLEHFTDSTDVISMRSVFSQT 240  
QY 263 HEWRNESAVLGCVFIFLLSTRYFSKPRPFVWSAMAPLTSVILGSLIAYTHAEKHGV 322  
DB 241 HEWRNESVILGCGFLFLLSTRYFSIKPKPFVWAMAPLTSVILGSLIAYTHAEKHGV 300  
QY 323 EVIGELKGLNPPSLTNLVFVSPYMTAVKTIIVGIIISLAEGIAVGRSFAMKYNIDG 382  
DB 301 QVIGDLKGLNPLSGDLIFTSPYMTAVKTIIVGIIISLAEGIAVGRSFAMKYNIDG 360  
QY 383 NKEMIAIGTMNVVGSFTSCYLTGTPFSRNVNAGCKTAASNIIMSLAVMLTLFLTP 442  
DB 361 NKEMIAFGMMNIVGSFTSCYLTGTPFSRNVNAGCKTAMSNIVMAIIVMTLLFLTP 420  
QY 443 FHYTFLVLSAIIIVSAMGLIDYEAHILFKVDKDFVVCMSAYIGVPGSVGIEIGVIAI 502  
DB 421 FHYTFLVLSAIIIVSAMGLIDYEAHILFKVDKDFVVCMSAYIGVPGSVGIEIGVIAI 480  
QY 503 VTSRVLVLLFIARPRFTVLGNIPNSVIVRNVEHYQNAKHVPGMLILEIDAPIYFANASYL 562  
DB 481 AISIARLLLFVSRPKTAVKGNIPNSVIVRNVEHYQNAKHVPGMLILEIDAPIYFANASYL 540  
QY 563 RERITRWIDEEERIKATGETSLQVVIDMSAVGNIDTSGISMLEEVKKITERRELQVL 622



```
Db 541 REIIIRWIDEEERVKQSGESSLOYIILDSAVGNIDTSGISMMVEIKKVIDRRALKVL 600
Qy 623 VNPVSEVMKLNKSKF-QNHHLGKKWLYLTVEEAVGACNFNLRASTNPKKDETEGNV 680
Db 601 SNPKGEVVKLTTSKFTGDHGLKEWMFLTVGEAVEACSYMLHTFKTEP-ASKNEPNV 658

RESULT 2
Q70GL1
ID Q70GL1 PRELIMINARY; PRT; 658 AA.
AC Q70GL1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Sulphate transporter.
GN Name=bst3.1;
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RA Buchner P., Stuijver E.E., Hawkesford M.J., de Kok L.J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ581745; CAB46442.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008271; F:sulfate porter activity; IEA.
DR GO; GO:0008272; P:sulfate transport; IEA.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transpt.
DR InterPro; IPR001902; Sulp_transpt.
DR Pfam; PF01740; STAS; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
DR PROSITE; PS50801; STAS; 1.
SQ SEQUENCE 658 AA; 72734 MW; 2EP39BD46340E17E CRC64;

Query Match 74.4%; Score 2596.5; DB 2; Length 658;
Best Local Similarity 75.0%; Pred. No. 2e-162;
Matches 494; Conservative 80; Mismatches 82; Indels 3; Gaps 3;

Qy 24 MGSVDVEYPLGMN-FERHVQVEVPPQPFKSLKSLKSETFPDDPLRQPKNKPASKFP 82
Db 1 MGTEDNTFPQGAEEPHRRHVAEPDQPFKSLQYSVKETLPDDPFPQFKNQTTSRQV 60
Qy 83 MLGLQFPFFIFEWAPKTYTFOFLKADLIAGITIASLAIPQGISYAKLANLPPIGLYSSFI 142
Db VLGLKTFPLPILWAPLYNFKLFSDDLIGITIASLAIPQGISYAKLANLPPIGLYSSFV 120
Qy 143 PPLIYAMGSSRDLAGVTGAVGSLMGSLMSNAVDPNEDPKLYLHLAFTATLTFAGVFOAA 202
Db PPLVFAVLGSSKDLAVGTGAVGSLLTGAMLSKEVDAEKDPKLYLHLAFTATLTFAGVLEAS 180
Qy 203 LGLFRGLIVDFLSHATITGFMGAATVCLQOLKSLGLEHPTHGADIISVNRVFTQT 262
Db LGLFRGLGVDFLSHATITGFMGAATVLSLOOLKGLFGLKHFTEATDVISNRSVSQT 240
Qy 263 HEWRWESAVLGCVFIPEFLLSSTRYFSKRPFFWVSAMAPLTSVILGSLLVYFTHSKGV 322
Db HQWRWESVGLGCCFLPFLLSSTRYFSKRPFFWVAAMAPLTSVILGSLLVYFTHSRHG 300
Qy 323 EVIGELKGLGNPPLSLNLPVSPYMTAVKTVGIWVGIIISLAEGIAVGRSFAMYNKYNIDG 382
Db QVIGDLKGLGNPLSVSDLVFTSPYMTALKTGIIITLAEGLAVGRSFAMFNKYNIDG 360
Qy 383 NKEMIAIGTNVVGSTFTSCVLTGTPFSRSANVTNAGCKTAASNIIMSLAWMLTLLPL 442
Db NKEMIAFGMNNIVGSEFTSCVLTGTPFSRSANVTNAGCKTAVSNIVMAIIMVMTLLFHTPF 420
Qy 443 FHYTPLVLSAIVSAMGLIIDEAAIHLFKVDKDFVVCMSAYIGVFGSVIEIGLVIAI 502
```

```
Db 421 FHYTPLVLSIIIMVAMGLIDIQAAILHLWKVDKDFVVCMSAYFGVFGSVIEIGLVAV 480
Qy 503 VTSVRVLLFIARPTFVLGNIPNSVIYRNVEHYQNAKHPVGMILLILEIDAPIYFANASYL 562
Db VISIARLLLFVSRPRTAVKGNIPNTMIYRNTDQPYSRIVPGLLILEIDAPIYFANAGYL 540
Qy 563 RERITRWIDEEERIKATGETSLOYVVIDMSAVGNIDTSGISMLSEVKKITERRELQVL 622
Db RERITRWIDEEEDRIKASGNSLQVVIDMSAVGNIDTSGISMMEEIKIMDRRELKVL 600
Qy 623 VNPVSEVMKLNKSKF-QNHHLGKKWLYLTVEEAVGACNFNLRASTNPKKDETEGNV 680
Db ANPKGEVVKLTTSKFTGEIGENLGEWMFLTVGEAVEACSPWLHTSKTEPASKE-EPN 658

RESULT 3
Q8H7X7
ID Q8H7X7 PRELIMINARY; PRT; 637 AA.
AC Q8H7X7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative sulfate transporter AtST1.
GN Name=OJ1607A12.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Sasaki C.,
RA Currie J., Collura K.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC105729; AAN06871.1; -.
DR Gramene; Q8H7X7; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008271; F:sulfate porter activity; IEA.
DR GO; GO:0008272; P:sulfate transport; IEA.
DR InterPro; IPR003658; Antisig_antgnt.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transpt.
DR InterPro; IPR001902; Sulp_transpt.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR ProDom; PD005210; Antisig_antgnt; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
DR PROSITE; PS50801; STAS; 1.
SQ SEQUENCE 637 AA; 69159 MW; DD09D98D86F35059 CRC64;

Query Match 66.9%; Score 2326.5; DB 2; Length 637;
Best Local Similarity 70.7%; Pred. No. 1.1e-144;
Matches 440; Conservative 82; Mismatches 99; Indels 1; Gaps 1;

Qy 46 VPPQPPFKSLKSLKSETFPDDPLRQF-KNKPASKKFMGLQFPFPIFEWAPKYTFQEL 104
Db 1 MPAAPFPLETGLGNMKETFLPDDPFRVVRERCGGRRAAAALRYVFPFWEAPSYLTGL 60
Qy 105 KADLAGITIASLAIPQGISYAKLANLPPIGLYSSFIPEPLIYAMGSSRDLAGVTGAVG 164
Db KSDLTAGITIASLAIPQGISYAKLANLPVILGLYSSFPVPLIYAMGSSRDLAGVTGAVA 120
Qy 165 SLLMGSLMSNAVDPNEDPKLYLHLAFTATLTFAGVFOAALGLFRGLIVDFLSHATIG 224
Db SLLIGSLMSEESAAEDPALYLHVALTATFFAGVFOALLGLVLRGLFVDFLSHATIG 180
Qy 225 GGAATVCLQOLKSLGLEHPTHGADIISVNRVFTQTHWRWESAVLGCVFIPEFLLS 284
Db GGAATVCLQOLKSLGLEHPTHGADIISVNRVFTQTHWRWESAVLGCVFIPEFLLS 240
Qy 285 YFSKKRPFFWVSAMAPLTSVILGSLLVYFTHAEKHGVEVIGELKGLNPPLSLNLFVS 344
Db YFSKKRPFFWVSAMAPLTSVILGSLLVYFTHAEKHGVEVIGELKGLNPPLSLNLFVS 300
```



```
FT TRANSMEM 391 411 Potential.
FT TRANSMEM 412 432 Potential.
FT DOMAIN 433 447 Extracellular (Potential).
FT TRANSMEM 448 468 Potential.
FT DOMAIN 469 646 Cytoplasmic (Potential).
FT DOMAIN 504 627 STAS.
SQ SEQUENCE 646 AA; 71269 MW; B611E17A9B036E3 CRC64;

Query Match
Best Local Similarity 66.6%; Score 2314; DB 1; Length 646;
Matches 430; Conservative 104; Mismatches 102; Indels 4; Gaps 2;

Qy 42 HOVEVPPQPPFKSLKYSKLETFFPPDPLQPKNK-PASKKFWLGLQFPPIPEWAPKYT 100
Db 10 HOVEIPPPQPLKSLKNTLNEILFADDPFRIRNESKTSKKIBGLRHFVPILEWARGVS 69

Qy 101 FOFLKADLIAGITIASLAIPQGISYAKLANLPILGLYSSFIPIPIYAMGSSRDLAAGT 160
Db 70 LEYLKSDVISGIIIASLAIPQGISYAKLANLPILGLYSSLVPLVYVAINGSSRDLAAGT 129

Qy 161 VAGSLMGLMNAVDNEDPKLYLHLATATLFAGVFOAALGLFRLGLIVDFLSHATI 220
Db 130 VAVASLLTAAMLGKEVNAVNPXLYLHLATATLFAAGLMQTCGLLRGLFVWEILSHAAI 189

Qy 221 IGFWGGAATVVCLOQLKSLIIGLEHFTGADIISVMRSVFTQTHEWRESAVLGCVPFEL 280
Db 190 VGFMGGAATVVCLOQLKGLGLHFTHTSTDIVTLRSIFSQSHWRWESVGLGCCFLIFL 249

Qy 281 LSTRYFSKKRPFWISAMAPLTSVLGSLVFTFAEKHGVEVIGELKKGLNPPSLTNL 340
Db 250 LTTKYISKKPKLFWISAMSPVSVFTGTFILFLHDQFGIOFIELKKGINPPSIHL 309

Qy 341 VFVSPYMTTAVKGTGI VVGIIISLAEGIAVGRSPAMYKNYNIDGNKEMIAIGTMNVGSFTS 400
Db 310 VPTPPYVMLAKVGIITVIALAEGIAVGRSPAMYKNYNIDGNKEMIAFGCMNLSGFS 369

Qy 401 CYLTGPFSSAVNAGCTAASNIIMSLAVMLTLFLTPLFHYPLVLSLIIIVSAML 460
Db 370 CYLTGPFSSAVNAGCTALSNVMAVAVTLLFLTPLFYPLVLSLIIIAAML 429

Qy 461 GLDYEAAIHLFKVDKDFDVVCMASAYIGVVGFSVEIGLVIATVISLRLFTARPTFV 520
Db 430 GLVDYEAAILHLKLDKDFDFVCLSAVIGVVGFGIEGLIISVGSVWRLVLFVGRKIYV 489

Qy 521 LGNI PNSVIYRNVEHYQNAKHVPCMLILEIDAPIYFANASYLRERITRWIDEBEERIKAT 580
Db 490 MGNIQNSEIYRNIEHYPOAITRSSLLIHLHDGPIYFANSTYLRDRIGRWIDEEDKLRTS 549

Qy 581 GETSLQYIVLDMASVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKLNKSKFQN 640
Db 550 GDLSLQYIVLDMASVGNIDTSGISMLBELNKLIGRELKLVIANPGAEVMKLSKSTFIE 609

Qy 641 HLKCKKIYILTVBEAVGACNFNRASKTNPKKDETEGNNV 680
Db 610 SIGKERIYLTVBEAVACDFMLHTAKPDSVPPE---FNNV 646

RESULT 5
Q70BL1
ID Q70BL1 PRELIMINARY; PRT; 650 AA.
AC Q70BL1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Sulfate transporter.
GN Name=1373.2;
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RA Buchner P., Stuver E.E., Hawkesford M.J., de Kok L.J.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ601439; CAE53112.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008271; F:sulfate porter activity; IEA.
DR GO; GO:0008272; P:sulfate transport; IEA.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transpt.
DR InterPro; IPR001902; Sulp_transpt.
DR Pfam; PF01740; STAS; 1.
DR TIGREMS; TIGR00815; sulp; 1.
DR PROSITE; P50801; STAS; 1.
SQ SEQUENCE 650 AA; 71703 MW; 3880636AF83017DA CRC64;

Query Match
Best Local Similarity 65.4%; Score 2274.5; DB 2; Length 650;
Matches 428; Conservative 105; Mismatches 108; Indels 13; Gaps 4;

Qy 30 EYPLGNFERNVHQVEVPPQPPFKSLKYSKLETFFPPDPLQPKNK-PASKKFWLGLQF 88
Db 7 QYP-----QVETPPQPPFLKSLKNTLNEILFADDPFGKIRNESKTSKKIBGLRH 56

Qy 89 FPIPEWAPKTYTQFLKADLIAGITIASLAIPQGISYAKLANLPILGLYSSFIPIPIYA 148
Db 57 VFPILEWAGYNLDYLSKSDVISGIIIASLAIPQGISYAKLANLPILGLYSSVFPVMA 116

Qy 149 MNGSSRDLAAGTAVAGSLMGLMNAVDNEDPKLYLHLATATLFAGVFOAALGLFRL 208
Db 117 IMGSSKDLAVGIVAVASLLTAAMLGKEVSAVENPKLYLHLATSTFFAGLMQTCGLRL 176

Qy 209 GLIVDFLSHATIIGRWGGAATVVCLOQLKSLIIGLEHFTGADIISVMRSVFTQTHEWRE 268
Db 177 GFLVEITLSHAAIIGFMAGAAATVVCLOQLKGLGLSHFTSTDVSVLSRIFSQSPVWR 236

Qy 269 SAVLGCVPFELLSSTYFSKKRPFWISAMAPLTSVLGSLVFTFAEKHGVEVIGEL 328
Db 237 SGLGCCFLFELPITKYISKRPKPLFWISAMAPLSVIFGSLFVYFLHAQFHGIQIIGEL 296

Qy 329 KXGLNPPSIITNLVFSVPMYTTAVKGTGI VVGIIISLAEGIAVGRSPAMYKNYNIDGNKEMIA 388
Db 297 EKGINPPSIITNLVFTSPYVMLAKTGIIITVGLAELAGIAVGRSPAMYKNYNIDGNKEMIA 356

Qy 389 IGTNVVGSFTCYLTGPFSSAVNAGCTAASNIIMSLAVMLTLFLTPLFHYTPL 448
Db 357 FGMNIFGSSCYLTGPFSSAVNAGCTAASNIIMSLAVMLTLFLTPLFHYTPL 416

Qy 449 VVLSAIIIVSAMLGLDYEAAIHLFKVDKDFDVVCMASAYIGVVGFSVEIGLVIATVISLRL 508
Db 417 VVLSIIIAAMLGLVDYEAAIHLFKVDKDFDVVCMASAYIGVVGFSVEIGLVIATVISLRL 476

Qy 509 VLLFTARPTFVLGNIPNSVIYRNVEHYQNAKHVPCMLILEIDAPIYFANASYLRERITR 568
Db 477 LLLFVGRPKIYVNGNIQNTPEYNTIEQYQATLSSIIILHDGPIYFANASYLRDRIGR 536

Qy 569 WIDEBEERIKATGETSLQYIVLDMASVGNIDTSGISMLBEVKKITERRELQVLVNPVSE 628
Db 537 WIDEBEERIKATGETSLQYIVLDMASVGNIDTSGISMLBEVKKITERRELQVLVNPVSE 596

Qy 629 VMKLNKSKFQNLKCKKIYILTVBEAVGACNFNRASK-TNPKKDETE-GWNNV 680
Db 597 LMKKLSKSKFPETIGKOWIHLTVAEAVSACDFVHLHTAKPDSPEKTSREPEFNNV 650

RESULT 6
Q7XE18
ID Q7XE18 PRELIMINARY; PRT; 658 AA.
AC Q7XE18;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to sulfate transporter (AtST1).
GN ORFNames=OSUNBa005313.6;
```



```
-----
CC      EMBL; AB023423; BAA75015.1; -.
DR      EMBL; AC002311; AAC00610.1; -.
DR      PIR; B86365; B86365.
DR      GenBank; 2362; 189.
DR      InterPro; IPR002645; STAS.
DR      InterPro; IPR001902; Sulph_transpt.
DR      Pfam; PF01740; STAS; 1.
DR      Pfam; PF00916; Sulfate_transp; 1.
DR      TIGRFAMs; TIGR00815; sulp; 1.
DR      PROSITE; PS01130; SLC26A; FALSE_NEG.
DR      PROSITE; PS0801; STAS; 1.
KW      Multigene family; Sulfate transport; Symport; Transmembrane;
KW      Transport.
FT      DOMAIN 1 69 Cytoplasmic (Potential).
FT      TRANSMEM 70 90 Potential.
FT      DOMAIN 91 92 Extracellular (Potential).
FT      TRANSMEM 93 113 Potential.
FT      DOMAIN 114 117 Cytoplasmic (Potential).
FT      TRANSMEM 118 138 Potential.
FT      DOMAIN 139 144 Extracellular (Potential).
FT      TRANSMEM 145 165 Potential.
FT      DOMAIN 166 171 Cytoplasmic (Potential).
FT      TRANSMEM 172 192 Potential.
FT      DOMAIN 193 223 Extracellular (Potential).
FT      TRANSMEM 224 244 Potential.
FT      DOMAIN 245 256 Cytoplasmic (Potential).
FT      TRANSMEM 257 277 Potential.
FT      DOMAIN 278 309 Extracellular (Potential).
FT      TRANSMEM 310 330 Potential.
FT      DOMAIN 331 347 Cytoplasmic (Potential).
FT      TRANSMEM 348 368 Potential.
FT      DOMAIN 369 384 Extracellular (Potential).
FT      TRANSMEM 385 405 Potential.
FT      DOMAIN 406 410 Cytoplasmic (Potential).
FT      TRANSMEM 411 431 Potential.
FT      DOMAIN 432 441 Extracellular (Potential).
FT      TRANSMEM 442 462 Potential.
FT      DOMAIN 463 631 Cytoplasmic (Potential).
FT      TRANSMEM 497 621 STAS.
FT      CONFLICT 89 90 KL -> NV (in Ref. 1).
SQ      SEQUENCE 631 AA; 65070 MW; 21E48F67125E6200 CRC64;

Query Match 54.6%; Score 1898.5; DB 1; Length 631;
Best Local Similarity 57.1%; Pred. No. 1.6e-116;
Matches 353; Conservative 125; Mismatches 139; Indels 1; Gaps 1;

QY 41 VHOVEVPPPOPFKSLKYSKLTETFPDDPLRQFNKPKASKKFMGLGQFFPIPEWAPKXT 100
DB 3 VHKVVAAPHKSTVAKLTKLTKETFPDDPLRQFNKPKASKKFMGLGQFFPIPEWAPKXT 62
QY 101 FQFLKADLIAGITIASLAIPQGISYAKLANLPILGLYSFIPPLIYAMGSSRDLAGVT 160
DB 63 FSLKSDVVSGLTIASLAIPQGISYAKLANLPILGLYSFIPPLIYAMGSSRDLAGVP 122
QY 161 VAVGSLMLGMSLNAVDNEDPKLYLHLATATLPAQVFOAALGLFRGLIVDFLGHATI 220
DB 123 VSTASILGSMRLQVSPVDDPVLFLQAFSTFFAGLFOASILGLRGLIIFDLSKATL 182
QY 221 IGFWGGATVVCLOQLKSIILGLEHFTHGADIISVMRSVFTQTHWESAVLGCVFIFEL 280
DB 183 IGFWGGAAITVLSQQLKGLIGITFTKHMSSVPLSVFQHTNWSQTTVMGVCFLFL 242
QY 281 LSTRYSKKRPFRFWVSAMAPLTSVILGLLVFTFAEKGHVRIIGELKGLNPPSLTNL 340
DB 243 LSTRHLMSKKPKPLFWVSAGAPLISVIVSTLLVFFRAERAGISVIGKLPGLNPPSNML 302
QY 341 VFVSPYMTTAVKGTIVVGIISLAEGIAVGRSFMKYKNYINDGNKEMIAIGTMNVGSFTS 400
DB 303 QFHGSHLALVAKTGLVTGIVSLTEGIAVGRFTAALKKNYHVDGNKEMIAIGLMNVGSATS 362
QY 401 CYLTGTPFSNAYNAGCTAASNIIMSLAVMLTLFLTPHYPPLVLSIIIVSAML 460
DB 401 CYLTGTPFSNAYNAGCTAASNIIMSLAVMLTLFLTPHYPPLVLSIIIVSAML 460
-----
```

```
DB 363 CYTTGAFSRSVANNAGAKTAVSNIVMSVTVMVTLFLMPLFEYTPNVVLGAIIVTAVI 422
QY 461 GLIDYEAAIHLFKVDKDFVVCMSAYIGVVGFSVEIGLVIATVISVLRLFLFIARPTFV 520
DB 423 GLIDLPAACHIKIDKDFELVMLCAFFGIFLSVQNGLAIAVGLSLFKILMQVTRPKMVI 482
QY 521 LGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLREIRLTRWIDEEERIKAT 580
DB 483 MGNIPGTDIYRDLHYKEAQRIPGFLVLSIESPVNFANSNYLTERTSRWIECEEBAQE 542
QY 581 GETSLQYIIDSAVGNIDTSGISMLEEVKKITERRELQVLVLPVSEVMKLNKSKFQ 640
DB 543 KHSSQLQFLILEMSAVSGVDVTNGVSPFKELKTKTKAKKDIELVFNPLSEVEVKLQRADEQ 602
QY 641 H-LGKKVIYLTVEEAVGA 657
DB 603 EFMRPFEFLTVAEAVAS 620

RESULT 8
Q9AT47 PRELIMINARY; PRT; 656 AA.
AC Q9AT47;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sulfate transporter 2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22983287; PubMed=129317983;
RX Howarth J R, Fourcroy P, Davidian J C., Smith F W., Hawkesford M J.;
RT "Cloning of two contrasting high-affinity sulfate transporters from
RT tomato induced by low sulfate and infection by the vascular pathogen
RT Verticillium dahliae.";
RL Planta 218:58-64 (2003).
DR EMBL; AF347614; AAK27688.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008271; F:sulfate porter activity; IEA.
DR GO; GO:0008272; P:sulfate transport; IEA.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR011547; Sulph_transpt.
DR InterPro; IPR001902; Sulp_transpt.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
DR PROSITE; PS01130; SLC26A; 1.
DR PROSITE; PS0801; STAS; 1.
SQ SEQUENCE 656 AA; 72556 MW; DE2A1DD2EF81C097 CRC64;

Query Match 54.1%; Score 1880.5; DB 2; Length 656;
Best Local Similarity 55.6%; Pred. No. 2.6e-115;
Matches 349; Conservative 119; Mismatches 159; Indels 1; Gaps 1;

QY 36 NNFERVHOVEVPPPOPFKSLKYSKLTETFPDDPLRQFNKPKASKKFMGLGQFFPIFEW 95
DB 25 NQTPYHKVGVPKQVFKETTKVETTFADDPPLRSFKQQRSLKVLGLQAFILDW 84
QY 96 AFKYPFQFLKADLIAGITIASLAIPQGISYAKLANLPILGLYSFIPPLIYAMGSSRD 155
DB 85 GRSYNLRKFRGDLISGLTIASLICIPQDIGYKLANLDPOGLYSSFPVPLVAFMGSSRD 144
QY 156 LAVGTAVGSLMLGMSLNAVDNEDPKLYLHLATATLPAQVFOAALGLFRGLIVDFL 215
DB 145 IAGPVAVVSLLLGTLRLNEIDPSKHPAEYLAFTATFAGITATLGLRGLIDFL 204
QY 216 SHATIIIGFGGATVVCLOQLKSIILGLEHFTHGADIISVMRSVF-TQTHWESAVLGC 274
DB 205 SHAAVVGFGGAATITALLQQLKGLIGITFTKHKADIIISVMKSVFHSVEHWNQTLIGA 264
```

```
QY 275 VFIFFLLSTRYFKRPRFFWVSAMAPLTSVILGSLVYTHAEKHGVEVIGELKKGLNP 334
D 265 TELTELLFAKYVGKKNKLLFWPAIAPLISVILSTFFVYTHADKRGVAIVGRIGKINP 324
QY 335 PSUTNLVFSVPYMTAVKTIIVGIIISLABGIAVGRFSFMYKNYNIDGNKEMIAITGMNV 394
D 325 PSVDKIYFSGDGLMKGIRTGIVAGMIALTEAVAIGRTFASMKDYQLDGNKEMVAGMNI 384
QY 395 VGSFTSCYLTGTPFSFSAVNNAGCKTAASNIIMSLAVMLTLLPLPHYTHPLVLSAI 454
D 385 VGSMTSCYATSFSAVNNAGCQAFSAFNNVSVVFTLFTPLFKFTFNALAAI 444
QY 455 IVSAMLGLIDYEAAILHFKVDKDFVVCMSAYILGVVFGSVEIGLVAIVSLRVLLETA 514
D 445 IISAVLGLIDYEAAILWKIDKDFVACIGAFGVVFSVEIGLLAVISPAKILLQVT 504
QY 515 RPTFVLGNTPNSVIVRNVHYQNAKHVPQWLILEIADIYFANASYLRRITRWIDEE 574
D 505 RPRATLGRIPRTNVYRNTQQYPEATKVPGLVIRVDSAIYFSNSVNIKERILRWLWDE 564
QY 575 ERUKATGETSLOQVVIDMSAVGNDTSGISLMEVKKIITERLOQLVIVNPVSEVMKLN 634
D 565 EQRKATSDPKIQFLIVEMSPVTDITSGHALBELHRSRIKRNQVLVLSNPGRWIDKUL 624
QY 635 KSKFQNLGKKWYLYLVEEAVGACNFNL 662
D 625 ASKEPQOIGEDKIFLTVADAVLTCSLKL 652

RESULT 9
ST34 ARATH
ID ST34 ARATH STANDARD; PRT; 653 AA.
AC Q9LM6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Probable sulfate transporter 3.4.
GN Names=SU1R3.4; OrderedLocustNames=At3g15990; ORFNames=MSL1.3;
OS Arabidopsis thaliana (Mouse-ear cress)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Takahashi H., Watanabe-Takahashi A., Saito K., Yamaya T.;
RT "cDNA for sulfate transporter Sultr3.4.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kaneo T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
CC -1- FUNCTION: H(+)/sulfate cotransporter that may play a role in the
CC regulation of sulfate assimilation (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the SLC26A/SulP transporter (TC 2.A.53)
CC family.
CC -1- SIMILARITY: Contains 1 STAS domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; AB054645; BAB21264.1; -.
DR EMBL; AB012247; BAB02665.1; -.
DR GeneFarm; 2697; 189.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR001902; Sulph_transpt.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
DR PROSITE; PS01130; SLC26A; FALSE_NEG.
DR PROSITE; PS0801; STAS; 1.
KW Multigene family; Sulfate transport; Symport; Transmembrane;
KW Transport.
FT DOMAIN 1 92 Cytoplasmic (Potential).
FT TRANSMEM 93 113 Potential.
FT DOMAIN 114 115 Extracellular (Potential).
FT TRANSMEM 116 136 Potential.
FT DOMAIN 137 140 Cytoplasmic (Potential).
FT TRANSMEM 141 161 Potential.
FT DOMAIN 162 167 Extracellular (Potential).
FT TRANSMEM 168 188 Potential.
FT DOMAIN 189 194 Cytoplasmic (Potential).
FT TRANSMEM 195 215 Potential.
FT DOMAIN 216 247 Extracellular (Potential).
FT TRANSMEM 248 268 Potential.
FT DOMAIN 269 279 Cytoplasmic (Potential).
FT TRANSMEM 280 300 Potential.
FT DOMAIN 301 331 Extracellular (Potential).
FT TRANSMEM 332 352 Potential.
FT DOMAIN 353 370 Cytoplasmic (Potential).
FT TRANSMEM 371 391 Potential.
FT DOMAIN 392 407 Extracellular (Potential).
FT TRANSMEM 408 428 Potential.
FT DOMAIN 429 433 Cytoplasmic (Potential).
FT TRANSMEM 434 454 Potential.
FT DOMAIN 455 471 Extracellular (Potential).
FT TRANSMEM 472 492 Potential.
FT DOMAIN 493 520 Cytoplasmic (Potential).
FT DOMAIN 520 643 STAS.
SQ SEQUENCE 653 AA; 71176 MW; 88FBF572E32E08BF CRC64;

Query Match 54.0%; Score 1875; DB 1; Length 653;
Best Local Similarity 55.5%; Pred. No. 5.9e-115;
Matches 361; Conservative 118; Mismatches 158; Indels 14; Gaps 2;

QY 21 IEDMGSDVDEYPLGMNN-----PFRHOVEVPPQPPFKSLKYSKTEFFDDPDLQRFK 74
D 8 VEDMAS-----PNNGTAGETVVEIHSVCLPPKTAFOKLKRVGDVFFDDPDLQRF 59
QY 75 NKPASKKFMGLQFPFPIFEWAPKYTFQFLKADLIAGITIASLAIPQGISYAKLANLPPI 134
D 60 NQTWERNVILQSLFPPIFTWGSQYDLKLRSDVISGLTIASLAIPQGISYAKLANLPPI 119
QY 135 LGLYSFPTPLIYAMWGSSRLAVGTAVGSLMGSMLSNAVDNEDPKLYLHLAFTATL 194
D 120 VGLYSFVPPLIYAVLGSRLHVA GPVSIA SLVMSLSSESPTSQDSLYLKLAFSTTF 179
QY 195 FAGVFAALGLFRLGLIVDFLSHATIGFMGGAATVCLQQLKSLILGLEHFTGADIISV 254
D 180 FAGVFAQSLGLRLGLFMDIFLSKATLIGTAGAIVSVLQQLKGLIGIVHFTGKQIVPV 239
QY 255 MRSVFTQTHWEWSAVLGCVFIFLLSTRYFSKRPFRFFWVSAMAPLTSVILGSLVYF 314
D 240 MSSVFNRHRSWETIVMGIGFLSILLTRHISMKPKLFWISAAASPLASVLIISLTVL 299
QY 315 THAEKHGVEIGELKKGLNPPSLTNLVFSPVMYTTAVKTIIVGIIISLABGIAVGRSFA 374
D 300 IRSKNTHAISFIIHGLPKGLNPPSLNMLYFSGAHLALAIKTIITGILSLTEGAVGRTPAS 359
QY 375 YKNYINIDGNKEMIAITGMNVVGSFSTSCYLTTPFSFSAVNNAGCKTAASNIIMSLAVML 434
D 360 LKNYQVNGKNEMALGFMNMGAGSCTSCYVTTGFSFSAVNNAGAKTAVSNIVMASAVLV 419
QY 435 TLLFLTPLFHYTHPLVLSAIIVSAMLGLIDYEAAILHFKVDKDFVVCMSAYILGVVFGSV 494
```







```

Db 213 GFAGAAITIALQOLGFLGIANFTKKSDIVSVKMGVHGHWNQITLIGATLAF 272
Qy 281 LSTREYSKKPRFFWWSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGLNPPSLTNL 340
Db 273 LVAKYIGKRNKKLFWWSAIPLTSVISTFTFVYITRADKHGVAIVKRNIRGINPPSASLT 332
Qy 341 VFVSPYMTTAVKGTGIVVGIISLAEGIAVGRSFAMYNKYNIDGNKEMIAIGTNNVGSFTS 400
Db 333 YFTGPIVATGFKIGIVAGMIGLTEAIGTFFAALDKYRIDGNKEMMALGTNNVGSLSIS 392
Qy 401 CYLTTGPFSSAVYNNAGCKTAASNTIMSLAVMLTLFLTPPLPHYTPLVVLSAIIIVSAML 460
Db 393 CYVATGFSRSVAVNMGCKTAVSNVMSVIMVMTLLITLTPPKYTPNALSSIIISAVL 452
Qy 461 GLIDYEAAHLFKVDKDFVWCMSAIVGVFGSVEIGLVIAIVISVLVLLFTARRPTFV 520
Db 453 GLIDYESAVLIWKDKLDFLACGAPFGVIFSSVEYGLLIAVAISLAKILLQVTRPTVL 512
Qy 521 LGNIPNSVIYRNVEHYONAKHVPCMLILEIDAPIYFANASYLRERTRMIDBEEBRIKAT 580
Db 513 LGNLPQTIYRNVEQYFDPATKVFGLIVRDSAIYFTNSYVYKERTLRMLRDEEBEQOQ 572
Qy 581 GETSLQYVIDMSAVGNIDPISGMSLEEVKKTITRERLQVLVNPVSEVMKKLNKSKFQN 640
Db 573 KLTKEFLVLDLSPVIDIDTSGIHALDELAKALEKRIQLVLTNPGPVIQKLSAKFTD 632
Qy 641 HLGKKWYLYTVEAV 655
Db 633 MIGEDNIFLTVGDAV 647

RESULT 12
ST12_ARATH
ID ST12_ARATH STANDARD; PRT; 653 AA.
AC G9MAX3; Q9SH00;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Sulfate transporter 1.2.
GN Name=SULTR1.2; OrderedLocustNames=A1g78000; ORFNames=F28K19.22;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND INDUCTION.
RC STRAIN=cv. Columbia;
RX MEDLINE=21836421; PubMed=11846879;
RA Yoshimoto N., Takahashi H., Smith F.W., Yamaya T., Saito K.;
RT "Two distinct high-affinity sulfate transporters with different
RT inducibilities mediate uptake of sulfate in Arabidopsis roots.;"
RL Plant J. 29:465-473(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altshuler H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.-L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaynsberg M., Vysotskaia V.S., Walker M.,

```

```

RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana";
RL Nature 408:816-820(2000).
RN [3]
RP FUNCTION, AND MUTANTS SEL1.
RX MEDLINE=21836422; PubMed=11846880;
RA Shibagaki N., Rose A., McDermott J.P., Fujiwara T., Hayashi H.,
RA Yoneyama T., Davies J.P.;
RT "Selenate-resistant mutants of Arabidopsis thaliana identify Sultr1.2,
RT a sulfate transporter required for efficient transport of sulfate into
RT roots.;"
RL Plant J. 29:475-486(2002).
CC -!- FUNCTION: High-affinity H(+)/sulfate cotransporter that mediates
CC the uptake of the environmental sulfate by plant roots. Plays a
CC central role in the regulation of sulfate assimilation.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Expressed in lateral root cap, root hairs,
CC epidermal and cortical cells of roots.
CC -!- INDUCTION: In roots by sulfate starvation.
CC -!- MISCELLANEOUS: Sel1 mutations in the gene lead to the resistance
CC of the plant to selenate, a toxic analog of sulfate. (TC 2.A.53)
CC -!- SIMILARITY: Belongs to the SLC26A/SulP transporter (TC 2.A.53)
CC family.
CC -!- SIMILARITY: Contains 1 STAS domain.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB042322; BAA95484.1; -.
CC EMBL; AC009243; AAF17685.1; ALT_SEQ.
CC PIR; D96809; D96809.
CC GeneFam; 2355; 189.
CC InterPro; IPR002645; STAS.
CC InterPro; IPR001902; Sulph_transpt.
CC Pfam; PF01740; STAS; 1.
CC Pfam; PF00916; Sulfate_transp; 1.
CC ProDom; PD006364; DUP52; 1.
CC TIGRFAMs; TIGR00815; sulp; 1.
CC PROSITE; PS01130; SLC26A; 1.
CC PROSITE; PS50801; STAS; 1.
CC KW Multigene family; Sulfate transport; Symport; Transmembrane;
CC Transport.
KW DOMAIN 1 91 Cytoplasmic (Potential).
FT TRANSMEM 92 112 Potential.
FT DOMAIN 113 116 Extracellular (Potential).
FT TRANSMEM 117 137 Potential.
FT DOMAIN 138 141 Cytoplasmic (Potential).
FT TRANSMEM 142 162 Potential.
FT DOMAIN 163 173 Extracellular (Potential).
FT TRANSMEM 174 194 Potential.
FT TRANSMEM 195 215 Potential.
FT DOMAIN 216 253 Extracellular (Potential).
FT TRANSMEM 254 274 Potential.
FT DOMAIN 275 280 Cytoplasmic (Potential).
FT TRANSMEM 281 301 Potential.
FT DOMAIN 302 339 Extracellular (Potential).
FT TRANSMEM 340 360 Potential.
FT DOMAIN 361 372 Cytoplasmic (Potential).
FT TRANSMEM 373 393 Potential.
FT DOMAIN 394 409 Extracellular (Potential).
FT TRANSMEM 410 430 Potential.
FT DOMAIN 431 438 Cytoplasmic (Potential).
FT TRANSMEM 439 459 Potential.
FT DOMAIN 460 466 Extracellular (Potential).
FT TRANSMEM 467 487 Potential.

```

FT	DOMAIN	488	653	Cytoplasmic (Potential).
FT	DOMAIN	522	645	STAS.
FT	MUTAGEN	511	511	1-57: In sell-8; reduces drastically sulfate transport activity.
SQ	SEQUENCE	653 AA; 71707 MW; 8019CADD09ADE46 CRC64;		
Query Match				
Best Local Similarity 53.0%; Score 1841.5; DB 1; Length 653;				
Matches 343; Conservative 117; Mismatches 157; Indels 1; Gaps 1;				
Qy	42	HOVEVPPQPPFKSLKYSKLETFPPDDPLRQFNKPKASKFMLGLQFPFFIFWAPKYTF	101	
Db	28	HKVGIPKQNWDFWYTKETFTFDDPLRDFDKQPKSKQFMLGLQSFVFFVDFGWNYYTF	87	
Qy	102	QFLKADLIAGITIASLAPGDISYAKLANLPPIGLYSSFIPPLIYVAMGSSRDLAGTV	161	
Db	88	KXFRGLISGLTTLASLIPQDIGYAKLANLPKLYSSFPVPLVACMGSSRDIAIGPV	147	
Qy	162	AVGSLMGSMNAVDNEDPKLYLHLAFTATLFAGVQALGLFRGLIVDFLSHATII	221	
Db	148	AVVSLGLTGLLRABIDPNTSPDEYLRLAFTATTFAGITEAALGFFRGLIDFLSHAUV	207	
Qy	222	GFMGAAITVLCLOQKSGILGLEHFTHGADIISVMSRVSFTQTHE-WRWESAVLGCVPFFFL	280	
Db	208	GFMGAAITVLCLOQKSGILGLEHFTHGADIISVMSRVSFTQTHE-WRWESAVLGCVPFFFL	267	
Qy	281	LSTRYFKKRPFPFWVSAMAPLTSVLGSLLVYFTHAEKHGVVEVIGELKKGLNPPSLTNL	340	
Db	268	LTSKIIGKSKGLFWPFAIAPLISVIVTFFVITRADKQGVQIVKHLDOGINPSSPHLI	327	
Qy	341	VFVSPYMTTAVKTVIGVIGIISLAEGIAVGRSPAMYKYNIDGNKEMIAIGTMNVVGSFTS	400	
Db	328	YFTGDNLAGIRIGWAGWALTEAVAIGRTFAAMKDYQDGNKEMVAGLMGMNVVGSMS	387	
Qy	401	CYLTGTPSRSVAVNAGCKTAASNIIMSLAVMLTLLPLFHYTPLVLSAIIIVSAML	460	
Db	388	CYVATGSFSRSVAVNAGCKTAVNSNIIMSVLTLTLLPLFYTPNAILAIIINAVI	447	
Qy	461	GLIDYEAAILHFKVDKDFVVCMSAYIGVVFGSVGEIGLVAIIVISVLRLVLLFTARPRTFV	520	
Db	448	PLIDIOQAAILIFKVDKLDFTACIAGPFGVIFVSVEIGLLIAVISFPAKILLQVTRPTAV	507	
Qy	521	LGNIPNSVIRNVEHYQNAKHVPCMLLEIDAPIFANASYLRERITRWIDEEERIKAT	580	
Db	508	LGNIPTRSVNRNTQQPEATWPGVLTVRVDSTAIYFSNSYVRERTQRWLHEEKEVKAA	567	
Qy	581	GETSLOVVIDMSAVGNIDTSGISMLREVKKITERRELOLVLPVNPVSEVMKLNKSKFON	640	
Db	568	SLPRIQPLIEMSPVTDIDSGIHALEDLYKSLQKRDIIQILANPGPLVIGKLHLSHAD	627	
Qy	641	HLGKKWITYLTVEEAVGAC	658	
Db	628	MLGQDNIYLTVADAVEAC	645	
RESULT 13				
Q7X610 PRELIMINARY; PRT; 629 AA.				
AC	Q7X610			
DT	01-OCT-2003	(TrEMBLrel. 25, Created)		
DT	01-OCT-2003	(TrEMBLrel. 25, Last sequence update)		
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)		
DE	OSJNBA0060D06.5	protein (OSJNBB0059K02.23 protein)		
GN	Name=OSJNBA0060D06.5; Synonyms=OSJNBB0059K02.23;			
OS	Oryza sativa (japonica cultivar-group)			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzoideae; Oryza.			
OX	NCBI_TaxID=39947;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	Pubmed=12447439; DOI=10.1038/nature01183;			
RA	Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,			

RA	Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,			
RA	Weng Q., Zhang L., Lu Y., Mu J., Sun T., Lei H., Li T., Hu H., Fan D.,			
RA	Liu X., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Fan D.,			
RA	Wu M., Zhang R., Zhou B., Chen Z., Jin Z., Wang R., Yin H.,			
RA	Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,			
RA	Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,			
RA	Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,			
RA	Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,			
RA	Han B.;			
RT	"Sequence and analysis of rice chromosome 4.";			
RL	Nature 420:316-320(2002).			
DR	EMBL; AL606690; CAE03539.2; -			
DR	EMBL; AL606692; CAE04513.2; -			
DR	Gramene; Q7X610; -			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0008271; P:sulfate porter activity; IEA.			
DR	GO; GO:0008272; P:sulfate transport; IEA.			
DR	InterPro; IPR002645; STAS.			
DR	InterPro; IPR011547; Sulph. transp.			
DR	InterPro; IPR001902; Sulp. transp.			
DR	Pfam; PF01740; STAS; 1.			
DR	Pfam; PF00916; Sulfate transp; 1.			
DR	TIGRFAMS; TIGR00815; sulP; 1.			
DR	PROSITE; PS50801; STAS; 1.			
SQ	SEQUENCE 629 AA; 68354 MW; 8629F81E11EE9FAD CRC64;			
Query Match 52.8%; Score 1836.5; DB 2; Length 629;				
Best Local Similarity 57.0%; Pred. No. 1.9e-112;				
Matches 352; Conservative 109; Mismatches 154; Indels 3; Gaps 2;				
Qy	41	VHQUEVPPQPPFKSLKYSKLETFPPDDPLRQFNKPKASKFMLGLQFPFFIFWAPKYT	100	
Db	4	LHKVAPPAQSTASKLAKARVKETFFPDPPRGFGKPLTTKVMVAVQYLPFLIDMWPSYS	63	
Qy	101	FQFLKADLIAGITIASLAPGDISYAKLANLPPIGLYSSFIPPLIYVAMGSSRDLAGTV	160	
Db	64	FSLFKSDDLAVGLTTLASLAPGDISYAKLANLPPIGLYSSFPVPLVACMGSSRDLAGTV	123	
Qy	161	VAVGSLMGSMNAVDNEDPKLYLHLAFTATLFAGVQALGLFRGLIVDFLSHATII	220	
Db	124	VSTASLIMGSMNAVDNEDPKLYLHLAFTATLFAGVQALGLFRGLIVDFLSHATII	183	
Qy	221	IGPMGAAITVLCLOQKSGILGLEHFTHGADIISVMSRVSFTQTHEWESAVLGCVPFFL	280	
Db	184	VGFMAGAAITVLCLOQKSGILGLEHFTHGADIISVMSRVSFTQTHEWESAVLGCVPFFL	243	
Qy	281	LSTRYFKKRPFPFWVSAMAPLTSVLGSLLVYFTHAEKHGVVEVIGELKKGLNPPSLTNL	340	
Db	244	LTAHVSMKWPKLFWVSACAPLACTIVSTLLVFLFPAQKHGIIIGLQKGLNRPWDKL	303	
Qy	341	VFVSPYMTTAVKTVIGVIGIISLAEGIAVGRSPAMYKYNIDGNKEMIAIGTMNVVGSFTS	400	
Db	304	LFDPQYLGTVTKGLVTGIIISLTEGVAVGRTFASLKDYVDGNKEMVAGLMNVGSCTS	363	
Qy	401	CYLTGTPSRSVAVNAGCKTAASNIIMSLAVMLTLLPLFHYTPLVLSAIIIVSAML	460	
Db	364	CYVTTGAFSRSVAVNAGCKTAMSNVIMALTVMVTLLFLMPLFVYTPNVVLAIIIAAVI	423	
Qy	461	GLIDYEAAILHFKVDKDFVVCMSAYIGVVFGSVGEIGLVAIIVISVLRLVLLFTARPRTFV	520	
Db	424	GLIDLPVAVNIMKQKDFLVCLCAFAGVIFISVQQGLAIVAGISIFRVLQLTRPKMMI	483	
Qy	521	LGNIPNSVIRNVEHYQNAKHVPCMLLEIDAPIFANASYLRERITRWIDEEERIKAT	580	
Db	484	QGNIKGTDIYRNHLHQKQARVPGFLITVEAFINPANTNYLNERIKRWI--EEESSAGT	541	
Qy	581	GETSLOVVIDMSAVGNIDTSGISMLREVKKITERRELOLVLPVNPVSEVMKLNKSK-FQ	639	
Db	542	KQSELHFVLDLSAVPAIDTSGISFLIDLKSTKHEGLLELVNPTGVEVMEKIQANDAH	601	
Qy	640	NHLGKKWITYLTVEEAVGA	657	
Db	602	GHPKDSLSLYLTGEEAVAS	619	





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 31, 2005, 09:41:38 ; Search time 971 Seconds

(without alignments)  
4145.648 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLLRNVI.....NLRSKTNPKDKETEGWNNV 680

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10762049/runat.30082005.121244.5960/app.query.fasta\_1.839  
-DB=N Geneseq 16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10762049 @CGN 1.1.470 @runat.30082005.121244.5960 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002as: \*  
7: Geneseqn2002bs: \*  
8: Geneseqn2003as: \*  
9: Geneseqn2003bs: \*  
10: Geneseqn2003cs: \*  
11: Geneseqn2003ds: \*  
12: Geneseqn2004as: \*  
13: Geneseqn2004bs: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3475	100.0	2449	3 AAZ50489	Aaz50489 Soybean s
2	2641.5	76.0	1977	3 AAC43310	Aac43310 Arabidops
3	2638.5	75.9	2112	3 AAC43093	Aac43093 Arabidops
4	2595.5	74.7	1977	6 ABZ13697	Abz13697 Arabidops
5	2595.5	74.7	1977	8 ADA68514	Ada68514 Arabidops

6	2330.5	67.1	1974	8 ADA70002	Ada70002 Rice gene
7	2184	62.3	1981	3 AAZ50482	Aaz50482 Corn sulp
8	1902.5	54.7	2006	3 AAC42229	Aac42229 Arabidops
9	1901	54.7	2067	3 AAZ50488	Aaz50488 Soybean s
10	1898.5	54.6	1896	8 ADA68554	Ada68554 Arabidops
11	1844.5	53.1	2279	3 AAZ50481	Aaz50481 Corn sulp
12	1817	52.3	1940	8 ADA71315	Ada71315 Rice gene
13	1807.5	52.0	1371	6 ABQ82719	Abq82719 Arabidops
14	1807.5	52.0	2000	8 ADA70110	Ada70110 Rice gene
15	1799	51.8	2311	3 AAZ50490	Aaz50490 Wheat sul
16	1777.5	51.2	1989	8 ADA70466	Ada70466 Rice gene
17	1683.5	48.4	1974	8 ADA70351	Ada70351 Rice gene
18	1675.5	48.2	1947	8 ADA69770	Ada69770 Rice gene
19	1644	47.3	1896	8 ADA69681	Ada69681 Rice gene
20	1513.5	43.6	1980	12 ADJ40066	Adj40066 Plant cDN
21	1428	41.1	1983	8 ADA69804	Ada69804 Rice gene
22	1002	28.8	2058	6 ABZ12912	Abz12912 Arabidops
23	1002	28.8	2058	8 ADA68540	Ada68540 Arabidops
24	963.5	27.7	2022	3 AAZ50491	Aaz50491 Wheat sul
25	731.5	21.1	780	3 AAZ50484	Aaz50484 Attichoke
26	706.5	20.3	2829	11 ADM56356	Adm56356 Frog SLC2
27	687.5	19.8	1791	11 ABD08637	Abd08637 Pseudomon
28	679.5	19.6	2612	11 ADM56273	Adm56273 Mouse SLC
29	679.5	19.6	2750	11 ADM56275	Adm56275 Mouse SLC
30	676.5	19.5	1767	11 ABD08580	Abd08580 Pseudomon
31	675.5	19.4	1240	3 AAZ50483	Aaz50483 Corn sulp
32	670	19.3	2279	11 ADM56358	Adm56358 Pig SLC36
33	668.5	19.2	2654	11 ADM56269	Adm56269 Human SLC
34	668	19.2	2589	13 ADQ87181	Adq87181 Human tum
35	668	19.2	2748	11 ADM56271	Adm56271 Human SLC
36	665.5	19.2	2580	3 AAD00611	Aad00611 Human mem
37	664.5	19.1	2245	13 ACN43603	Acn43603 Human dia
38	662	19.1	2593	4 AAH75189	Aah75189 Nucleotid
39	662	19.1	12348	6 ABQ82327	Abq82327 Human NOV
40	662	19.0	13935	5 AAS74583	Aas74583 DNA encod
41	659	19.0	2241	13 ACN43604	Acn43604 Human dia
42	658.5	18.9	2241	3 AAZ29466	Aaz29466 P. chryso
43	657	18.9	2506	8 ABX12016	Abx12016 Transpore
44	651	18.7	2441	11 ADM57788	Adm57788 Murine pr
45	651	18.7	2441	12 ADE86099	Ade86099 cDNA enco

ALIGNMENTS

RESULT 1  
AAZ50489  
ID AAZ50489 standard; cDNA; 2449 BP.  
XX  
AC AAZ50489;  
XX  
DT 23-MAY-2000 (first entry)  
DE  
DE Soybean sulphate permease-2 cDNA clone.  
XX  
KW Sulphate Permease; sulphate assimilation protein; soybean; probe;  
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;  
KW antibody; screen; ss.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
CDS 2..2044  
FT /\*tag= a  
FT /product= "Soybean sulphate permease-2"  
FT /note= "Derived from clone sf11.pk0043.g10"  
XX  
FN WO200004154-A2.  
XX  
PD 27-JAN-2000.  
XX  
PF 13-JUL-1999; 99WO-US015810.  
XX  
PR 14-JUL-1998; 98US-0092833P.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX PA Allen SM, Falco SC, Thorpe CJ;  
 XX PI WPI; 2000-195025/17.  
 XX DR P-PSDB; AAY44943.  
 XX PT Nucleic acid fragments encoding sulfate assimilation proteins in plants  
 PT and seeds useful as probes for isolating cDNAs and genes encoding  
 PT homologous proteins, in producing transgenic plants.  
 XX PS Claim 3; Page 52-53; 79pp; English.  
 XX CC The present sequence is a cDNA clone encoding soybean sulphate permease,  
 CC a sulphate assimilation protein. This sequence is obtained from sfl1  
 CC library, clone sfl1.pk0043.g10, derived from soybean immature flower.  
 CC This sequence is used as a probe to isolate other plant sulphate  
 CC assimilation proteins, for genetic and physical mapping of related genes  
 CC and as markers of traits linked to the gene. This is useful for plant  
 CC breeding and to construct chimeric genes, used to create transgenic  
 CC plants with altered levels of sulphate permease. The sulphate permease  
 CC peptides are useful for producing antibodies, that are used to screen and  
 CC isolate cDNA clones  
 XX SQ Sequence 2449 BP; 719 A; 470 C; 543 G; 717 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0 Length: 2449  
 Score: 3475.00 Matches: 680  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0

US-10-762-049-18 (1-680) x AA250489 (1-2449)

QY 1 HisGluLeuAlaArgThrLeuSerTyrIleThrHisIleCysLeuLeuArgAsnThrIle 20  
 DB 2 CAGGAGCTAGCTGGACATTAAGTTATATATACACATATTGCTTGTGTAGAAATACTATT 61  
 QY 21 IleGluAspMetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArg 40  
 DB 62 ATTGAAGATATGGGAGTGTAGATTATGATGATACCTTTGGGCATGAACAACTTTGAGAGA 121  
 QY 41 ValHisGluValGluValProProGlnProPhePhePheLysSerLeuLysTyrSerIle 60  
 DB 122 GTGCACCAAGTCGAGGTTCACCGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTTG 181  
 QY 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80  
 DB 182 AAGGAGACTTTCTTCCCTGATGACCTTTGAGGCAGTTCAAGAACCAAGCCAGCTTCCAAG 241  
 QY 81 LysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThr 100  
 DB 242 AAGTTTCATGCTGGCCCTTCAGTCTTCTTCCCAATTTTGGGATGGGCTCCCAATACACC 301  
 QY 101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120  
 DB 302 TTTTCAGTTCTTTGAAAGCTGACCTTCATGCTGGCATCACCATGCTAGCTTGGCCATTCCT 361  
 QY 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSer 140  
 DB 362 CAGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCCAATTTCTTGGACTATATTCGAGC 421  
 QY 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160  
 DB 422 TTTATACCAACCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481  
 QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180  
 DB 482 GTGGCCGGTTGGATCGCTTCTGATGGGTTTCAATGTTGAGTAAATGCCGTTGATCCCAATGAA 541

QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200  
 DB 542 GACCCAAAGCTTTACTCCACCTGGCTTTCACAGCTACATTATTTGCTGGTGTTCAG 601  
 QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220  
 DB 602 GCTGCCCTTGGGTCTGTATTAGGTGGGGTTGATCGTGGATTCTGTGTACATGCAACCAT 661  
 QY 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240  
 DB 662 ATAGGGTTTCATGGGAGGAGCAGCCAGGTGGTGTGTCTGCAGCAACTAAAAATCGATTCT 721  
 QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260  
 DB 722 GGCTTTGAGCAATTCACCCATGGAGCTGATATCATATCAGTATGGCTCTGTTCAC 781  
 QY 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280  
 DB 782 CAAACTCATGATGGAGGTGGGAAAGTGTGTGTAGGATGTCTCTTCATATTTCTTCCTC 841  
 QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300  
 DB 842 CTTAGCACAAAGATACTTCAGCAAAAACACCAAGGTTTTTTTGGGTGTACGAATGGCG 901  
 QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320  
 DB 902 CCATTGACGTCCGTTATATTGGGAAGTCTCTTGGTTTTATTCTCACTCACGCCGAGAAGC 961  
 QY 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340  
 DB 962 GGTGTGTAAGTCATAGGAGAACTGAAGAAGGGTTTTGAATCCACCATCACTCAAAATCTG 1021  
 QY 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360  
 DB 1022 GTATTGTGTCCCTTACATGACTACAGCTGTCAAAACCTGGCATTTGCTGTGGCATCA 1081  
 QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380  
 DB 1082 TCACTTGGCGAAGGAATAGCAGTAGGAGAACTTTGCAATGTATAAAAAATTACAATATT 1141  
 QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400  
 DB 1142 GATGGCAACAAAGAGATGATAGTATTGGGACCATGAACGTAGTTGGTTCTTTCACTCT 1201  
 QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420  
 DB 1202 TGCTACCTCACAAACAGGACCAATTTCCGCTTCGGCTGTGAACATATAACCTGGATGCAAG 1261  
 QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440  
 DB 1262 ACAGCAGCTTCCAACTATATATATGTCACCTTGCAAGTAATGTTGACATTGTTATTCCTG 1321  
 QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValIleSerAlaMetLeu 460  
 DB 1322 CCCTTGTTCATATACATCCCTCGTGGTGTCTATCAGCTATTATCGTATCTCCAATGCTTT 1381  
 QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480  
 DB 1382 GGACTCATAGATTATGAGCAGCCATCCATCTATTATTAAGTTGACAAATTTGACCTTTGG 1441  
 QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500  
 DB 1442 GTGTGCATGAGTGCATACATGGCGTGTCTTTGGCAGTGTGTGAAATTTGGCTTAGTCATA 1501  
 QY 501 AlaIleValIleSerValLeuArgValIleLeuPheIleAlaArgProArgThrPheVal 520  
 DB 1502 GCTATTGTAAATATCTGACTTCGGGTACTTCTATTATTATGCAAGGCCAAGGACATTTCGT 1561  
 QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540  
 DB 1562 TTGGGCAACATTTCCAAATTCGTGATATACCGAAATTTTGACCACTATCAAAATGC AAA 1621  
 QY 541 HisValProGlyMetLeuLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560



Db	1622	CCTGGAATGCTAAATCTAGAGATTGATGCCAAATTTACTTTGCCAATGCCAGC	PR	14-MAY-1999;	99US-0134219P.
			PR	14-MAY-1999;	99US-0134221P.
Qy	561	TyrLeuArgGluArgIleThrArgTrrPileAspGluGluGluArgIleLysAlaThr	PR	14-MAY-1999;	99US-0134370P.
			PR	18-MAY-1999;	99US-0134768P.
Db	1682	TAITTAAGAGAAAGGATCACAGGTGGATTGATGAAGAAGAAAGAAATTAAGCTACA	PR	19-MAY-1999;	99US-0134941P.
			PR	20-MAY-1999;	99US-0135124P.
Qy	581	GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAenIleAspThr	PR	21-MAY-1999;	99US-0135353P.
			PR	24-MAY-1999;	99US-0135629P.
Db	1742	GGGGAGACTAGTTTGCAGTATGTTATAATTGATATGAGTCTGTTGGAAACATTGTATACA	PR	25-MAY-1999;	99US-0136021P.
			PR	27-MAY-1999;	99US-0136392P.
Qy	601	SerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeu	PR	28-MAY-1999;	99US-0136782P.
			PR	01-JUN-1999;	99US-0137222P.
Db	1802	AGTGAATAAGTATGCTTGAAGAGGTGAAGAATTACAGAGAGAGAGCTACAGCTT	PR	03-JUN-1999;	99US-0137528P.
			PR	04-JUN-1999;	99US-0137502P.
Qy	621	ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn	PR	08-JUN-1999;	99US-0138094P.
			PR	10-JUN-1999;	99US-0138540P.
Db	1862	GTTTTGGTCAATCCTGTAACTGAAGTGAAGAACTGAACAAATCGAAAGTTCCAAAT	PR	14-JUN-1999;	99US-0138847P.
			PR	10-JUN-1999;	99US-0139119P.
Qy	641	HisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPhe	PR	16-JUN-1999;	99US-0139452P.
			PR	16-JUN-1999;	99US-0139453P.
Db	1922	CATTTAGGGAAAGAAATGGATCTATCTGACTGTTGAAGAGGCCGTTGGAGCATGCAACTTC	PR	17-JUN-1999;	99US-0139492P.
			PR	18-JUN-1999;	99US-0139454P.
Qy	661	AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal	PR	18-JUN-1999;	99US-0139455P.
			PR	18-JUN-1999;	99US-0139456P.
Db	1982	AACTACGTCAAGCAAAACCAACCCCAAGAAAGATGAACACAGAGGTTGGAAACAATGTG	PR	18-JUN-1999;	99US-0139457P.
			PR	18-JUN-1999;	99US-0139458P.
RESULT 2					
AAC43310					
ID	AAC43310	standard; DNA; 1977 BP.			
XX	AAC43310;				
XX	DT	17-OCT-2000 (first entry)			
XX	DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 38799.			
XX	XX				
KW	Hybridisation assay; genetic mapping; gene expression control;				
KW	protein identification; signal transduction pathway; metabolic pathway;				
KW	promoter; termination sequence; ss.				
XX	OS	Arabidopsis thaliana.			
XX	XX				
PN	EP1033405-A2.				
XX	XX				
PD	06-SEP-2000.				
XX	XX				
PF	25-FEB-2000; 2000EP-00301439.				
XX	XX				
PR	25-FEB-1999;	99US-0121825P.			
PR	05-MAR-1999;	99US-0123180P.			
PR	09-MAR-1999;	99US-0123548P.			
PR	23-MAR-1999;	99US-0125788P.			
PR	25-MAR-1999;	99US-0126264P.			
PR	23-MAR-1999;	99US-0126785P.			
PR	01-APR-1999;	99US-0127462P.			
PR	06-APR-1999;	99US-0128234P.			
PR	08-APR-1999;	99US-0128714P.			
PR	16-APR-1999;	99US-0129845P.			
PR	19-APR-1999;	99US-0130077P.			
PR	21-APR-1999;	99US-0130449P.			
PR	23-APR-1999;	99US-0130510P.			
PR	23-APR-1999;	99US-0130891P.			
PR	28-APR-1999;	99US-0131449P.			
PR	30-APR-1999;	99US-0132048P.			
PR	04-MAY-1999;	99US-0132407P.			
PR	05-MAY-1999;	99US-0132484P.			
PR	06-MAY-1999;	99US-0132485P.			
PR	06-MAY-1999;	99US-0132486P.			
PR	07-MAY-1999;	99US-0132487P.			
PR	11-MAY-1999;	99US-0112863P.			
PR	14-MAY-1999;	99US-0134256P.			
		99US-0134218P.			



QY 283 ThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeu 302  
Db 781 ACCAGATATTTCAGCATCAAGAAACCAAAATTCCTTTGGGTGGCGGATGGCTCTCTTTG 840  
QY 303 ThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVal 322  
Db 841 ACCTCAGTGATTCTTGGAGTCTCTTGGTTTACTTCACTCACGCTGAGAGACATGGTGT 900  
QY 323 GluValIleGlyLysLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPhe 342  
Db 901 CAAGTGATAGGACCTGGAAGAAGGGTTGAATCCACTCTCCGGTCTGATCTCATCTTT 960  
QY 343 ValSerProTyrMetThrThrAlaValLysThrGlyValValGlyLysLeuSerLeu 362  
Db 961 ACTTCCCTTACATGTCACAGCTGTCAAACTGGCCCTCATCTGCGCATCATGTCTC 1020  
QY 363 AlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGly 382  
Db 1021 GCTGAAGGAGTACAGTGGGAGGAGTTTGGATGTTCAAGAACTACAACATAGACGG 1080  
QY 383 AsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr 402  
Db 1081 AACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCTTCCATCTCTGTAC 1140  
QY 403 LeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAla 422  
Db 1141 CTCACAACCGGACCATTTTCAAGTCGGCAGTGAACACTACAACCGGGTTGCAAGACCGCA 1200  
QY 423 AlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuProLeu 442  
Db 1201 ATGTCCACATAGTATGGGATGGGTATGTTTACACTCTCTCTCTCACACCGCTT 1260  
QY 443 PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu 462  
Db 1261 TTTCACTACACCACTCGTCGTCCTCTCTGCCATCATCATATCCCAATGCTCGGACTC 1320  
QY 463 IleAspTyrClnuAlaAlaIleHisLeuPheLysValAspLysPheAspValValCys 482  
Db 1321 ATTGACTATCAAGCTGCCATCCATCTCTGGAAGTTGACAAGTTCCACTTCCTCGTCTGC 1380  
QY 483 MetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle 502  
Db 1381 ATGAGCGCTACCTTGGGTCTGTTATTCGGCAGTGTAGAGATTGGACTCTGCTGAGCGGTG 1440  
QY 503 ValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGly 522  
Db 1441 GCGATATCTATAGCAGGTGTTGCTGTTGTGTCGAGGCCAAAACACTGCGGTGAAGGGA 1500  
QY 523 AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVal 542  
Db 1501 AACATACCAACAGCATGATCTATAGAACACTGAGCAGTACCATCATCAAGAACCGTT 1560  
QY 543 ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu 562  
Db 1561 CTGTGTTATCTCATCTTGGAGATTGATGCTCCCATCTACTTGTCTAATGCCAGTTACTTG 1620  
QY 563 ArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGlu 582  
Db 1621 CGTGAAGAATCATAGGTGGATTGATGAAGAGAGAGAGAGTTAAACAATCAGGAGAG 1680  
QY 583 ThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly 602  
Db 1681 AGCAGCTTCAAAATATATTACTCGATGATGTCAGTGTGTTGGTAAATATCGACACACCGGT 1740  
QY 603 IleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValLeu 622  
Db 1741 ATTAGCATGATGTGGAATATTAGAAGTCTTTCACAGGAGACCGTTAAAGTTGGTATTG 1800  
QY 623 ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641  
Db 1801 TCAAAATCCAAAGAGAGGTGCGTGAAGAAATTAACCCAGATCCAAATTCATCGGTGATCAT 1860  
QY 642 LeuGlyLysLysTrpIleTyrLeuThrValGluAlaValGlyAlaCysAsnPheAsn 661

Db 1861 TTTGGGCAAGAGTGGATGTTCTTAACGGTAGGAGAACGAGCTTGTAGCTACATG 1920  
QY 662 LeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680  
Db 1921 CTTCAACAGCTTTAAACACCGAACCG---GCCTCCAAAACGAGCCTTTGGAACACGTA 1974  
RESULT 3  
AAC43093  
ID AAC43093 standard; DNA; 2112 BP.  
XX AAC43093;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 38002.  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-01411287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145911P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157533P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 22-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:  
Pred. NO.: 3.57e-262 Length: 2112  
Score: 2638.50 Matches: 499  
Percent Similarity: 88.16% Conservative: 82  
Best Local Similarity: 75.72% Mismatches: 75  
Query Match: 75.93% Indels: 3  
DB: 3 Gaps: 3

US-10-762-049-18 (1-680) x AAC43093 (1-2112)









[illegible]

Db	1860	TTTGGGCAAGAGTGGATGTTCTTAACGGTAGGAGAAGCAGTGGAGGCTTTGTAGCTACAT	1919
Qy	661	nLeuAraGAlaSerLysThrAsnProLysLysAspGluThrGluGlyTirAsnAsnVal	680
Db	1920	GCTTCACACGTTTAAACCGAACCG---GCCTCAAAACGAGCCTTGGACACGTA	1974
RESULT 6			
ADA70002	ID	ADA70002 standard; DNA; 1974 BP.	
XX	AC	ADA70002;	
XX	DT	20-NOV-2003 (first entry)	
XX	XX	Rice gene, SEQ ID 3325.	
XX	XX	Plant; bacterial infection; fungal infection; viral infection; rice;	
XX	XX	gene; ds.	
XX	OS	Oryza sativa.	
XX	PN	WO2003000898-A1.	
XX	XX	03-JAN-2003.	
XX	XX	22-JUN-2001; 2001WO-IB001105.	
XX	XX	22-JUN-2001; 2001WO-IB001105.	
XX	PA	(SYGN ) SYNGENTA PARTICIPATIONS AG.	
XX	PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;	
XX	PI	Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;	
XX	XX	WPI; 2003-175290/17.	
XX	PT	Identifying at least one gene involved in plant resistance or response to	
XX	PT	pathogenic infection for conferring resistance or tolerance to a plant	
XX	PT	bacterial, fungal or viral infection by determining or detecting plant	
XX	XX	gene expression.	
XX	PS	Claim 6; SEQ ID NO 3325; 899pp; English.	
XX	CC	The present invention relates to a method (M1) for identifying genes	
XX	CC	involved in plant resistance or response to pathogenic infection. M1	
XX	CC	comprises identifying a gene whose expression is significantly altered in	
XX	CC	the incompatible interaction of plant gene expression relative to	
XX	CC	expression of the gene in an uninfected plant, in a mutant plant that	
XX	CC	does not express a gene associated with response to pathogenic infection,	
XX	CC	or in a corresponding incompatible or compatible interaction. (M1) is	
XX	CC	useful for conferring resistance to resistance or tolerance to a plant to	
XX	CC	bacterial, fungal or viral infection. The present sequence was used to	
XX	CC	illustrate the invention.	
XX	XX		
XX	XX	Sequence 1974 BP; 335 A; 602 C; 657 G; 380 T; 0 U; 0 Other;	
Alignment Scores:			
XX	Pred. No.:	2,31e-230	Length: 1974
XX	Score:	2330.50	Matches: 441
XX	Percent Similarity:	83.84%	Conservative: 83
XX	Best Local Similarity:	70.56%	Mismatches: 100
XX	Query Match:	67.06%	Indels: 1
XX	DB:		Gaps: 1
XX	DB:		
US-10-762-049-18 (1-680) x ADA70002 (1-1974)			
Qy	43	GlnValGluValProProGlnProPheLysSerLeuLysTyrSerLeuLysGlu	62
Db	52	AGSGTCCGATGCGCGCGGAAGCGCTTCTGGAGACGCTGGGGGGAACATGAAGGAG	111
Qy	63	ThrPhePheProAspProLeuArgGlnPhe---LysAsnLysProAlaSerLysLys	81
Db	112	ACATTCCTCGGACACACCGCTTACGGGTGGTGGCGGGAGCGCGGGTGGCGGGCGC	171

```
QY 82 PheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPhe 101
Db |||:|||||
172 GCGCGCGCGCGCTCCGGTAGCTGCTCCGTTTCATGGAGTGGCGCGCTGATCACCTC 231
QY 102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121
Db |||:|||||
232 GGCACCTCAAGTCGACCTCATCGCGCATCACCATTCGCAGCTCGCCATCCCCAG 291
QY 122 GlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPhe 141
Db |||:|||||
292 GGCATCAGCTACGCCAAGCTCGCCAACTCCCTCCGCTCCTCGGCTCTATTCAGCTTC 351
QY 142 IleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal 161
Db |||:|||||
352 GTGCGCGCGCTGGGTAGCGCATGATGGGAGCTCGAGGGACCTGGCGGTGGGACCGTG 411
QY 162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db |||:|||||
412 GCGTGGCGTGGCTGCTGATCGGTTCGATGCTGAGCGAGGAGTGTGCGCGCGGAGGAC 471
QY 182 ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db |||:|||||
472 CCGCGCGCTGATCTGCACGCTCGCGCTCACGCCACCTTCTTCGCGCGGTTCACGGCG 531
QY 202 AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221
Db |||:|||||
532 CTGCTGGGGGCTTGGAGCTGGGGTTCATGTGACCTTCCTGCGCAGCCACCATGCTC 591
QY 222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db |||:|||||
592 GGGTTCATGGCGCGCGCCACCGCTGGTGGCTGCGCAGCAGCTCAAGGGCATGTTCCGC 651
QY 242 LeuGluHisPheThrHisGlyAlaAspIleLeuSerValMetArgSerValPheThrGln 261
Db |||:|||||
652 CTCGACCACCTCACCCACCGCCACCGACTCTGCTCTCCGTCATGAGCTCCGTTCTCCCAA 711
QY 262 ThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeuLeu 281
Db |||:|||||
712 ACCACCTCTGGCGATGGGAGAGGCTGCTCATGGGCTGCGGCTTCCTCTTCCTCCTC 771
QY 282 SerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaPro 301
Db |||:|||||
772 ATCACCGCTCTTCACGACGAGGAGGCGAAGGTCTCTGGGTATCTGCAGCTGCGCCA 831
QY 302 LeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGly 321
Db |||:|||||
832 TTGGCGTCTGTCATCATCGGAGGCTCTCTGTGTACCTCACACATGCTGAAACCATGGC 891
QY 322 ValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuVal 341
Db |||:|||||
892 ATTCAGGTGATTGTTTACCTGAAGAAGGGCTGAACCCCGCTGTCACCAAGCTGAAC 951
QY 342 PheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSer 361
Db |||:|||||
952 TTCTCTGCGCGTACATGATGCTGCGCTCGAAGACCGGATCATCACTGGCGTATTGCC 1011
QY 362 LeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAsp 381
Db |||:|||||
1012 CTCGCTGAGGGGATTGCGGTAGGAAGGAGCTTTTCAATGTTCAAGAATTACCATCGAC 1071
QY 382 GlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCys 401
Db |||:|||||
1072 GGAACAAGGAGATGATCGGGTTTGGAAACGATGAACATTGTGGGATCACTCACCTCTTGC 1131
QY 402 TyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThr 421
Db |||:|||||
1132 TACCTCACCCCGGTGCTTCTCGAGGTGGCGGTCACTACACCGCGGTGCAAGACG 1191
QY 422 AlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrPro 441
Db |||:|||||
1192 GCGATGTGGAACGTGATCATGCTGGTGGGTGATGATCACGCTGCTGTTCTTGACCGC 1251
```

```
QY 442 LeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGly 461
Db |||:|||||
1252 CTGTTCCACTACACCGCGCTGGTGGTCTCTCGCGCATCATGTGCGGATGCTCGGG 1311
QY 462 LeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValVal 481
Db |||:|||||
1312 CTGATCGACTACCGCGCGCGCTCCACCTGTGCGAGGTGCACAAGGTGGACTTCTCGGTC 1371
QY 482 CysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAla 501
Db |||:|||||
1372 TGCTCTCGGCGCTACCTCGGCTCGCTTTCGCGAGGCTCGAGATCGGCTCTCGTCGCC 1431
QY 502 IleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeu 521
Db |||:|||||
1432 GTCGGGATCTTCATCTCTCGCGTCTGCTGTGTCGCGCGCGGAGACGACGCTGCTC 1491
QY 522 GlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHis 541
Db |||:|||||
1492 GGGAAACATCCCAACACGATGATCTACCGCGGATGGACCATACACCGCGCGCAGAG 1551
QY 542 ValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyr 561
Db |||:|||||
1552 GTCCCGCGCTGCTCGTCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1611
QY 562 LeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGly 581
Db |||:|||||
1612 CTGCGTGGAGGATCGCGGTGGATCGACGACGAGGAGCAGTGTCAAGGAGAGGGC 1671
QY 582 GluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSer 601
Db |||:|||||
1672 GAGATGGCATTCAGTACGTCGTCCTCGACATGGTGGCTCGGAAGCATCGACACCG 1731
QY 602 GlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuVal 621
Db |||:|||||
1732 GGGACGAGCATGCTGGATGAATCAGAGAACCTCGACAGGAGGGGCTTCAGATTGTG 1791
QY 622 LeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHis 641
Db |||:|||||
1792 TTGGGAACCGGGGAGGAGATCATGAAGAAGTTGGACAGCTCCCAAGGTGCTTGAGCG 1851
QY 642 LeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661
Db |||:|||||
1852 ATCGGCGCATGATGGATCTTCCCAAGTGGGCGGCGGTGGCGGAGTGGCTCGTCTGTG 1911
QY 662 LeuArgAlaSerLys 666
Db |||:|||||
1912 ATGCACTCGCAGAAG 1926

RESULT 7
AAZ50482
ID AAZ50482 standard; cDNA; 1981 BP.
XX
AC AAZ50482;
XX
DT 23-MAY-2000 (first entry)
XX
DE Corn sulphate permease-2 cDNA clone.
XX
KW Sulphate Permease; sulphate assimilation protein; corn; contig; probe;
mapping; marker; plant breeding; chimeric gene; transgenic plant;
antibody; screen; ss.
XX
OS Zea mays.
XX
FH Key
CDS 2..1741
FT /*tag= a
FT /product= "Corn sulphate permease-2"
FT /note= "Derived from contig composed of clones
crln.pk0015.a2, p0006.cbvys25rb, p0072.comhc25r,
p0091.cmard29r and p0092.chwat43r"
XX
PN WO200004154-A2.
```





PR	06-AUG-1999;	99US-0147303P.	Score:	1902.50	Matches:	354
PR	06-AUG-1999;	99US-0147416P.	Percent Similarity:	76.89%	Conservative:	125
PR	09-AUG-1999;	99US-0147493P.	Best Local Similarity:	56.82%	Mismatches:	143
PR	09-AUG-1999;	99US-0147935P.	Query Match:	54.75%	Indels:	1
PR	10-AUG-1999;	99US-0148171P.	DB:	3	Gaps:	1
PR	11-AUG-1999;	99US-0148319P.				
PR	12-AUG-1999;	99US-0148341P.				
PR	13-AUG-1999;	99US-0148565P.				
PR	13-AUG-1999;	99US-0148684P.				
PR	16-AUG-1999;	99US-0149368P.				
PR	17-AUG-1999;	99US-0149175P.				
PR	18-AUG-1999;	99US-0149426P.				
PR	20-AUG-1999;	99US-0149723P.				
PR	20-AUG-1999;	99US-0149929P.				
PR	23-AUG-1999;	99US-0149902P.				
PR	23-AUG-1999;	99US-0149930P.				
PR	23-AUG-1999;	99US-0150566P.				
PR	25-AUG-1999;	99US-0150884P.				
PR	27-AUG-1999;	99US-0151065P.				
PR	27-AUG-1999;	99US-0151066P.				
PR	27-AUG-1999;	99US-0151080P.				
PR	30-AUG-1999;	99US-0151130P.				
PR	31-AUG-1999;	99US-0151438P.				
PR	01-SEP-1999;	99US-0151930P.				
PR	07-SEP-1999;	99US-0152363P.				
PR	10-SEP-1999;	99US-0153070P.				
PR	13-SEP-1999;	99US-0153758P.				
PR	15-SEP-1999;	99US-0154018P.				
PR	16-SEP-1999;	99US-0154039P.				
PR	20-SEP-1999;	99US-0154779P.				
PR	22-SEP-1999;	99US-0155139P.				
PR	23-SEP-1999;	99US-0155486P.				
PR	24-SEP-1999;	99US-0155659P.				
PR	28-SEP-1999;	99US-0156458P.				
PR	29-SEP-1999;	99US-0156596P.				
PR	04-OCT-1999;	99US-0157117P.				
PR	05-OCT-1999;	99US-0157753P.				
PR	06-OCT-1999;	99US-0157865P.				
PR	07-OCT-1999;	99US-0158029P.				
PR	08-OCT-1999;	99US-0158232P.				
PR	12-OCT-1999;	99US-0158369P.				
PR	13-OCT-1999;	99US-0159293P.				
PR	13-OCT-1999;	99US-0159294P.				
PR	13-OCT-1999;	99US-0159295P.				
PR	14-OCT-1999;	99US-0159329P.				
PR	14-OCT-1999;	99US-0159330P.				
PR	14-OCT-1999;	99US-0159331P.				
PR	14-OCT-1999;	99US-0159637P.				
PR	14-OCT-1999;	99US-0159638P.				
PR	18-OCT-1999;	99US-0159584P.				
PR	21-OCT-1999;	99US-0160741P.				
PR	21-OCT-1999;	99US-0160767P.				
PR	21-OCT-1999;	99US-0160768P.				
PR	21-OCT-1999;	99US-0160770P.				
PR	21-OCT-1999;	99US-0160814P.				
PR	21-OCT-1999;	99US-0160815P.				
PR	22-OCT-1999;	99US-0160980P.				
PR	22-OCT-1999;	99US-0160981P.				
PR	23-OCT-1999;	99US-0160989P.				
PR	25-OCT-1999;	99US-0161404P.				
PR	25-OCT-1999;	99US-0161405P.				
PR	25-OCT-1999;	99US-0161406P.				
PR	26-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.				
PR	26-OCT-1999;	99US-0161361P.				
PR	28-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161992P.				
PR	28-OCT-1999;	99US-0161993P.				
PR	29-OCT-1999;	99US-0162142P.				
PR	06-AUG-1999;	99US-0147303P.	Score:	1902.50	Matches:	354
PR	06-AUG-					

```
Db 1062 GTACCGGAATCGTCTCCCTCAGGAAGGAATCGAGTGGGAAGACATTTGCGAGCGCTA 1121
Qy LyeAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValVal 395
Db 1122 AAGAAGCTACACGCTAGTAGGAACAAGAGATGATGCCATTGGTCTGATGAACGTAGTA 1181
Qy GlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyr 415
Db 1182 GGCCTCGCCACATCTTGCTACGTACACACCGGAGCATTTCTTAGATCAGCGGGTCAAC 1241
Qy AsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThr 435
Db 1242 AAGCGGGAGCTAAACCGGAGCTTTCAACATTTGTTATGTCGCTACTGTTATGGTTACG 1301
Qy LeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIle 455
Db 1302 CTTCTCTCTTAATGCGGCTTTTGAATACACTCCCAATGTGCTCGTGGTGCATCAAT 1361
Qy ValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAsp 475
Db 1362 GTACCGCGGTCAATGGTCTCATCGACCTTCCCGCGGCTGTCCACATATGGAAGATCGAT 1421
Qy LysPheAspPheValValCysMetSerAlaTyrIleGlyValValPheGlySerValGlu 495
Db 1422 AATTTGATTTCTTGGTGATGCTTTCGCGGTTCTTGGTGTCATTTCTTCTCGTTCAA 1481
Qy IleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArg 515
Db 1482 AACGGTCTAGCCATTGCGTGGGGCTATCGTTGTTCAAGATATTTGATGCAAGTAAACA 1541
Qy ProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHis 535
Db 1542 CCGAAATGGTTTAATGGGTAATATTTCTTGGAAACGGATATATACCGAGATCTTCATCAT 1601
Qy TyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyr 555
Db 1602 TACAAGAACCAAGAGTCCCGGATTTCTTGTGTTAAGCATCGAATCTCTGTCAAT 1661
Qy PheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluGlu 575
Db 1662 TTCGCCAATTTCAACTACCTCACTGAAAGAACATCTCGTTGGATTGAAGATCGGAAGA 1721
Qy ArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaVal 595
Db 1722 GAGGAAGCTCAGAGAAGCATTTAGCTTACAGTTCTTGAATTTCTGAAATGTCCAGCGTG 1781
Qy GlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLysValIleThrGluArg 615
Db 1782 AGCGGTGTAGACACAACCGAGTGTCTTTTAAAGGAAGCTTAAAGAAACAACCGCCAG 1841
Qy ArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLys 635
Db 1842 AAGGACATCGAGCTGTGTTGTTGTAACCTCTAAGCGAAGTGTGGAGAAGCTTCAAGA 1901
Qy SerLysPheGlnAsnHis---LeuGlyLysLysTrpIleTyrLeuThrValGluGluAla 654
Db 1902 GCTGAGCAACAAAGAGTTTCATGAGCGCGGAGTTTCTTCTTAAACCGTGTGAGGCC 1961
Qy ValGlyAla 657
Db 1962 GTTGGGTGCG 1970

RESULT 9
AAZ50488
ID AAZ50488 standard; cDNA; 2067 BP.
XX
AC AAZ50488;
XX
DT 23-MAY-2000 (first entry)
XX
DE Soybean sulphate permease-1 cDNA clone.
XX
```

```
KW Sulphate Permease; sulphate assimilation protein; soybean; probe;
KW mapping; marker; plant breeding; chimeric gene; transgenic plant;
KW antibody; screen; ss.
XX
OS Glycine max.
XX
FH Location/Qualifiers
FT 2..1867
FT /*tag= a
FT /product= "Soybean sulphate permease-1"
FT /notes "Derived from clone ses2w.pk0031.b3"
XX
XX WO200004154-A2.
XX
XX 27-JAN-2000.
XX
XX 13-JUL-1999; 99WO-US015810.
XX
XX 14-JUL-1998; 98US-0092833P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Falco SC, Thorpe CJ;
XX
XX WPI: 2000-195025/17.
XX
XX P-PSDB; AAY44942.
XX
XX Nucleic acid fragments encoding sulfate assimilation proteins in plants
XX and seeds useful as probes for isolating cDNAs and genes encoding
XX homologous proteins, in producing transgenic plants.
XX
XX Claim 3; Page 49-50; 79pp; English.
XX
XX The present sequence is a cDNA clone encoding soybean sulphate permease,
XX a sulphate assimilation protein. This sequence is obtained from ses2w
XX library, clone ses2w.pk0031.b3, derived from soybean embryogenic
XX suspension 2 weeks after subculture. This sequence is used as a probe to
XX isolate other plant sulphate assimilation proteins, for genetic and
XX physical mapping of related genes and as markers of traits linked to the
XX gene. This is useful for plant breeding and to construct chimeric genes,
XX used to create transgenic plants with altered levels of sulphate
XX permease. The sulphate permease peptides are useful for producing
XX antibodies, that are used to screen and isolate cDNA clones
XX
XX SQ Sequence 2067 BP; 583 A; 450 C; 393 G; 641 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6-52e-186 Length: 2067
Score: 1901.00 Matches: 359
Percent Similarity: 76.48% Conservative: 106
Best Local Similarity: 59.05% Mismatches: 143
Query Match: 54.71% Indels: 0
DB: 3 Gaps: 0

US-10-762-049-18 (1-680) x AAZ50488 (1-2067)
Qy 48 ProGlnProPhePheLysSerLeuLysTyrSerLeuLysGluThrPhePheProAsp 67
Db 8 CCACACGACGACCACTCCACAACTTAGGCACAGAGTCTCCGAAATCTTCTCCAGAT 67
Qy 68 AspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPheMetLeuGlyLeuGln 87
Db 68 GACCCCTCTCCACCGTTCAGAACCAACTCGCTTTAANAAGTTCCTCTCGCAGCTCAG 127
Qy 88 PhePhePheProIlePheGluTrpAlaProLysTyrThrPheGlnPheLeuLysAlaAsp 107
Db 128 TATCTCTTCCCATTTTCGACTGGGCCCCAACTACAACTTACCCTTCTCCGCTCTGAC 187
Qy 108 LeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLys 127
Db 188 CTCATCTCTGGCTCACCATTGCCAGCTCGCATTCCTCAGGAATCAGTTATGCCAAG 247
Qy 128 LeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIleProLeuIleTyr 147
```

Sulphate Permease; sulphate assimilation protein; soybean; probe; mapping; marker; plant breeding; chimeric gene; transgenic plant; antibody; screen; ss.

Glycine max.

Location/Qualifiers  
2..1867  
/\*tag= a  
/product= "Soybean sulphate permease-1"  
/notes "Derived from clone ses2w.pk0031.b3"

WO200004154-A2.

27-JAN-2000.

13-JUL-1999; 99WO-US015810.

14-JUL-1998; 98US-0092833P.

(DUPO ) DU PONT DE NEMOURS & CO E I.

Allen SM, Falco SC, Thorpe CJ;

WPI: 2000-195025/17.

P-PSDB; AAY44942.

Nucleic acid fragments encoding sulfate assimilation proteins in plants and seeds useful as probes for isolating cDNAs and genes encoding homologous proteins, in producing transgenic plants.

Claim 3; Page 49-50; 79pp; English.

The present sequence is a cDNA clone encoding soybean sulphate permease, a sulphate assimilation protein. This sequence is obtained from ses2w library, clone ses2w.pk0031.b3, derived from soybean embryogenic suspension 2 weeks after subculture. This sequence is used as a probe to isolate other plant sulphate assimilation proteins, for genetic and physical mapping of related genes and as markers of traits linked to the gene. This is useful for plant breeding and to construct chimeric genes, used to create transgenic plants with altered levels of sulphate permease. The sulphate permease peptides are useful for producing antibodies, that are used to screen and isolate cDNA clones

Sequence 2067 BP; 583 A; 450 C; 393 G; 641 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6-52e-186 Length: 2067  
Score: 1901.00 Matches: 359  
Percent Similarity: 76.48% Conservative: 106  
Best Local Similarity: 59.05% Mismatches: 143  
Query Match: 54.71% Indels: 0  
DB: 3 Gaps: 0

US-10-762-049-18 (1-680) x AAZ50488 (1-2067)

Qy 48 ProGlnProPhePheLysSerLeuLysTyrSerLeuLysGluThrPhePheProAsp 67  
Db 8 CCACACGACGACCACTCCACAACTTAGGCACAGAGTCTCCGAAATCTTCTCCAGAT 67  
Qy 68 AspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPheMetLeuGlyLeuGln 87  
Db 68 GACCCCTCTCCACCGTTCAGAACCAACTCGCTTTAANAAGTTCCTCTCGCAGCTCAG 127  
Qy 88 PhePhePheProIlePheGluTrpAlaProLysTyrThrPheGlnPheLeuLysAlaAsp 107  
Db 128 TATCTCTTCCCATTTTCGACTGGGCCCCAACTACAACTTACCCTTCTCCGCTCTGAC 187  
Qy 108 LeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLys 127  
Db 188 CTCATCTCTGGCTCACCATTGCCAGCTCGCATTCCTCAGGAATCAGTTATGCCAAG 247  
Qy 128 LeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIleProLeuIleTyr 147

[illegible]

Db	1328	GGGGTCTGTTCATTTCACAGTCGCTTTAGGTCCTTGGGTATAGCGGTTATCATATATCATGTCCTC	1388
Qy	508	ArgValLeuLeuPheilleAlaAaRgProArGThrPheValLeuGlyAsnilleProAnSer	527
Db	1388	AAGATCCTGCTTCATGTCACCTGACCAACACATTTTGGTTTGGGAATATACACAGGAACA	1447
Qy	528	ValIleTyrArgAsnValGluHiisTyrGlnAsnAlaLysHisValProGlyMetLeulle	547
Db	1448	CAAAATATTCACAAACATATAACCAATACAAAAAGCTTTTAAGAGTTCCTTCATTTCTCAAT	1507
Qy	548	LeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrIleuArgGluArgilleThr	567
Db	1508	TTGGCTGTGGAGTCTCAATCTATTTTGTCTAACTCAACTTAATCTTCAAGAAAGGATACG	1567
Qy	568	ArgTIpIleAspGluGluGluArgilleLysAlaThrGlyGluThrSerLeuGlnTyr	587
Db	1568	AGATGGGTTTCGAAAGAGGAGACATATAAAGCTTAATAATGGAGCTCCATTTGAAGTGC	1627
Qy	588	ValIleIleAspMetSerAlaValGlyAsnilleAspThrSerGlyIleSerMetLeuGlu	607
Db	1628	ATAATTTTAGACATGACAGCTGTGCACAGGCACACAGACACAAGTGGGCTTGACACTTTATGT	1687
Qy	608	GluValLysLysIleThrGluArgGluGluGlnLeuValLeuValAsnProValSer	627
Db	1688	GAACTTAGAAGATGCTGGAGAGAGAGATCACTTGAGTTTGTGCTGGCAATCCGTGTGA	1747
Qy	628	GluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTIpIle	647
Db	1748	AATGTGATGAAATTTGCATTAAGTCAACATTTTCGATTTCTTTCGATTAAGAGATC	1807
Qy	648	TyrLeuThrValGluGluAlaVal	655
Db	1808	TATCTCACAGTGGGAGAGCTGTG	1831
RESULT 10			
ADA68554 standard; DNA; 1896 BP.			
AC	ADA68554;		
XX			
DT	20-NOV-2003 (first entry)		
XX			
DE	Arabidopsis thaliana gene, SEQ ID 528.		
KW	Plant; bacterial infection; fungal infection; viral infection; ds.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	WO2003000898-A1.		
XX			
PD	03-JAN-2003.		
XX			
PF	22-JUN-2001; 2001WO-IB001105.		
XX			
PR	22-JUN-2001; 2001WO-IB001105.		
XX			
PA	(SYGN ) SYNGENTA PARTICIPATIONS AG.		
PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;		
PI	Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;		
XX			
DR	WPI; 2003-175290/17.		
XX			
PT	Identifying at least one gene involved in plant resistance or response to		
PT	pathogenic infection for conferring resistance or tolerance to a plant		
PT	bacterial, fungal or viral infection by determining or detecting plant		
PT	gene expression.		
XX			
PS	Claim 6; SEQ ID NO 528; 899pp; English.		
XX			
CC	The present invention relates to a method (M1) for identifying genes		
CC	involved in plant resistance or response to pathogenic infection. M1		
CC	comprises identifying a gene whose expression is significantly altered in		



CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (MI) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.

XX  
SQ Sequence 1896 BP; 483 A; 465 C; 426 G; 522 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.04e-185 Length: 1896  
Score: 1898.50 Matches: 353  
Percent Similarity: 77.35% Conservative: 125  
Best Local Similarity: 57.12% Mismatches: 139  
Query Match: 54.63% Indels: 1  
DB: 8 Gaps: 1

US-10-762-049-18 (1-680) x ADA68554 (1-1896)

Qy 41 ValHisGlnValGluValProProGlnProPhePhePheLysSerLeuLysTyrSerLeu 60  
Db 7 GTGCACAAAGGTGGTGGTCCCGCGCATAGAGCAGCGGTGGCAAGCTTAAAGACAAAACTG 66  
Qy 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80  
Db 67 AAGGAGACTTCTTCCCGGACGATCCCTTAAGACAGTTTCAGAGACCAACCAACCGTACC 126  
Qy 81 LysPheMetLeuGlyLeuGlnPhePhePheProLysPheGluTrpAlaProLysTyrThr 100  
Db 127 AAGCTCATACGCGCGCTCAATACATTTTCCATCTCCATGGTGTCTGTAGTACAGC 186  
Qy 101 PheGlnPheLysAlaAspLeuAlaGlyLeuAlaGlyLeuAlaSerLeuAlaLeuPro 120  
Db 187 TTCAGCCTACTCAAAATCCGACGCTGTTTACAGTCTTACCATCGTAGTTTATTCCT 246  
Qy 121 GlnGlyLeuSerTyrAlaLysLeuAlaAsnLeuProProLysLeuGlyLeuTyrSerSer 140  
Db 247 CAGGGGATAGTTACGCGAAGCTAGCAAAATTTGCCAACATCTGTTGTCTATATCTGAGC 306  
Qy 141 PheileProLeuLeuTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160  
Db 307 TTCTGCCACCGTGGTTATGGCTTGGGAGCTTCAAGATCTAGAGCTGGGACCA 366  
Qy 161 ValAlaValGlySerLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180  
Db 367 GTCTCCATAGCTGCTTAATCTTAGGATCCATGAGGCAACAGTATCTCCCGTAGAC 426  
Qy 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200  
Db 427 GATCCTGTTCTTCTTACAGCTAGCCCTCTCTTACCTTCTTGTGCTCTTTCAA 486  
Qy 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuValAspPheLeuSerHisAlaThrIle 220  
Db 487 GCCTCTTGGAAATCTCAGCTGGGATTTATATAGATTTCTTCAAAAGGCGAGCTA 546  
Qy 221 IleGlyPheMetGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240  
Db 547 ATAGGGTTTATGGTGAGCAGCATAATGTATCTCACTCCACAGCTAAGGGTCTGCT 606  
Qy 241 GlyLeuGlnHisPheThrHisGlyAlaAspIleSerValMetArgSerValPheThr 260  
Db 607 GGGATACTCATTTTCAAAAGCATATAGTGTAGTCCCTTCTCTCTCTCTGTTTCCAA 666  
Qy 261 GlnThrHisGlnTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280  
Db 667 CACACCAAGGAGTGGTCATGGCAACATTTGTATGGAGTTTGTCTCTCTCTCTCTCT 726  
Qy 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300  
Db 727 CTCTCGACACGTCACTCAGCATGAGAGCCGAGCTGTTTGGGTCTCAGCGGAGCA 786  
Qy 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320

RESULT 11  
AAZ50481

Db 787 CCACCTTCTTCCGTTATCGTCTCTACACTTCTTGTCTTTTTCAGAGCGGCGTCAC 846  
Qy 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProSerLeuThrAsnLeu 340  
Db 847 GGAATCAGCGTCATCGGAAATTTACCAAGAGGTTTGAATCCACCGTCTTGGACATGCTT 906  
Qy 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIle 360  
Db 907 CAGTTTCCAGGTAGTCATCTCGCACTCGTCCCAAAACCGGACTCGTCACGGAATCGTC 966  
Qy 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380  
Db 967 TCCCTCAGGAAAGAAATCGAGTGGGAAGAAATTTGCAGCGTAAAGAAATACCACTA 1026  
Qy 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400  
Db 1027 GATGGAACAAAGAGATGATCGCATTTGGTCTGATGACGATAGGCTCTGCCACATCT 1086  
Qy 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420  
Db 1087 TGCTAGCTCACAAACCGGAGCAATTTCTAGATCAGCGGTCAACCAACACGCGGAGCTAAA 1146  
Qy 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440  
Db 1147 ACCGCAAGTTTCAACCAATTTGTTATGTCGTCACCTGTTATGTTACGCTTCTTCTCTAATG 1206  
Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460  
Db 1207 CCGCTTTTGAATACACTCCCAATGGTCTCGGTGCCATCATTTGTACCGGGTCAAT 1266  
Qy 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480  
Db 1267 GGTCTCATCGACCTTCCCGCGGCTGTCAATATGGAAGATCGATAAATTTGATTTCTTG 1326  
Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500  
Db 1327 GTGATCCTTTCGCGGTCTTTCGTTGTCATTTTCTTCCATTCGTTCAAAACCGTCTAGCCAT 1386  
Qy 501 AlaIleValIleSerValLeuValLeuPheIleAlaArgProArgThrPheVal 520  
Db 1387 CGCGTGGGCTATCGTTGTTCAAGATTTGATGCAAGTAAACAGGCGGAAATGGTTATA 1446  
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAlaLys 540  
Db 1447 ATGGGTATATCTCGAAGGATATATACCGAGATCTTCATCATTTACAAAGAAGACAA 1506  
Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560  
Db 1507 AGATCCCGGATTTCTTGTAAAGCATCGAATCTCTCTCAATTTCCCAATTTCTAAC 1566  
Qy 561 TyrLeuArgGluArgIleThrArgTTrpIleAspGluGluGluArgIleLysAlaThr 580  
Db 1567 TACCTCAGTAAAGAACATCTCTGTGGATTTGAAGATTCGAGAGAGGAGGAGCTCAAGAG 1626  
Qy 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600  
Db 1627 AAGCATTTCTAGCTACAGTCTTCTGATTTCTTGAATGTCAAGCTGAGCGGTGAGACACA 1686  
Qy 601 SerGlyIleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeu 620  
Db 1687 AACGGAGTCTCTTTTTTAAGGAACCTTAAGAAAAACACCCCAAGGAAGGACATCGAGCTT 1746  
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640  
Db 1747 GTGTTTGTGAACCTCTAAGCGAGGTGGAGAGCTTCAAGAGCTGACGACAAACAAA 1806  
Qy 641 His---LeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAla 657  
Db 1807 GAGTTTCATGAGCCCGAGTTTCTTCTTAAACGCTCGCTGAGGCGCTGCGTCG 1860



```
Db 1329 ACTGCCATGTCACACGATCATGGCGTCACTGTGATGGTCAGCTGCTTCCTCATG 1388
Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db 1389 CCACTGTTCTGTGTACACACCAACGTTGTCTCGAGCGATCATCATCGCGGTGATC 1448
Qy 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db 1449 GGCTGTATCATTTTCCCGGGGTGTACACATCTGGAAGATGACACAGATGGATTTTCTG 1508
Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db 1509 GTGTGCGTTTGGCGTTTGGCGGTCTTCTTCATCTCACTCAAGAGCCCTTGGCAT 1568
Qy 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
Db 1569 GCGTGTGTATATCTATATTATAGGTGTGTGATGCAGATCAACAGGCCGAAGATGATGTT 1628
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db 1629 CAAGGGAACATCAAGGGAGTGTATTTACAGAGACCTGCTCACTACAAAGGAGGCCAA 1688
Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db 1689 AGAGTTTCTGGTCTTGTATCTTGGCCATTGAAGCACCAGATAAATCTGCGCAATCCAC 1748
Qy 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
Db 1749 TACTGTAATGAAGATTAAAGATGATA-----GAGGAGAATCTTTTGAACAGGAT 1802
Qy 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db 1803 AAACATATCAATCCATTTTCAATCTGATCTGTGATCTGTCTGCAATTTGACACA 1862
Qy 601 SerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeu 620
Db 1863 AGTGGCATAGCGTTCCTCATGTACATAAAGAAATCAATAGAGAAACGTGTCTGGAGCTT 1922
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLys---SerLysPheGln 639
Db 1923 GTGCTGTCAATCAACTGGGAGTCAATGAGAAATACACGTCGAACAGGCTGAA 1982
Qy 640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 659
Db 1983 AACTATTTTAGGCGCAGATTGCTGTATCTGACCACTGGCGAAGCAATC-----GCTTCA 2036
Qy 660 PheAsnLeuArgAlaSerLysThrAsnProLysLysAsp 672
Db 2037 CTTTCTGCACTTGCCAAAGATGACAAACCCCTAAATGGAT 2075

RESULT 12
ADA71315
ID ADA71315 standard; DNA; 1940 BP.
XX
AC ADA71315;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 4638.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN W0203000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
```

```
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 6; SEQ ID NO 4638; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 1940 BP; 535 A; 421 C; 432 G; 552 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,89e-177 Length: 1940
Score: 1817.00 Matches: 341
Percent Similarity: 72.21% Conservative: 119
Best Local Similarity: 53.53% Mismatches: 163
Query Match: 52.29% Indels: 14
DB: 8 Gaps: 3

US-10-762-049-18 (1-680) x ADA71315 (1-1940)
Qy 34 GlyMetAsnAsnPheGluArgValHisGlnValProProGlnProPhePhe 53
Db 67 GGATCATCTCAGAAACCGGTGTGTACAGGTGAGGTCTCTCTCAAAAGAACTTTGCG 126
Qy 54 LysSerLeuLysTyrSerLeuLysGluThrPhePheProAspAspProLeuArgGlnPhe 73
Db 127 AGAGAGTTTAGAGACACATTTGAGAGAGACTTCTTCCATGACAAACCATTTACCCAGTAC 186
Qy 74 LysAsnLysProAlaSerLysLysPheMetLeuGlyLysGlnPhePheProIlePhe 93
Db 187 AAGGACCAATCTGGATCGGCAAGTTCAATGATGCGCGCTGAGTTCTGTTCCTCCATCTTT 246
Qy 94 GluTrpAlaProLysTyrThrPheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThr 113
Db 247 GAGTGGGCGAGGTGCTACAACTTAAGAAAGTTCAAGAGTGATTTAATTCGTGATTGACC 306
Qy 114 IleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAlaLeuPro 133
Db 307 ATTGCAAGTCTCTGCATTCCTCCAGGATATCGGGTATTTCGAAGCTTGTAACTTGGATGCA 366
Qy 134 IleLeuGlyLeuTyrSerPheIleProLeuIleTyrAlaMetMetGlySerSer 153
Db 367 CAATATGGAATATATCTCCAGCTTTGTTCCTCCCATATATATGCTGCAATGGGAGCTCA 426
Qy 154 ArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMetGlySerMetLeuSer 173
Db 427 AAGGATATTGCAATCGTTCGGTTGCAATG----- 456
Qy 174 AsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeuAlaPheThrAlaThr 193
Db 457 AATGAAGTCGACCTGTCTCAAAACACAGAGGAGGAATATCTGCGCTTCTGTTTACGGTACC 516
Qy 194 LeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuLeuValAsp 213
Db 517 TTCCTTTGGGTATCACTCAAGCAGCCCTAGGATTTTAAAGGCTAGGATTCCTTCATCGAG 576
```

```
QY 214 PheLeuSerHisAlaThrIleLeuGlyPheMetGlyGlyAlaAlaThrValValCysLeu 233
DB 577 TTCTTGTCATGCTGCAATGTTGGTTATGGAGGACGCCATTACTATTGCCCTT 636
QY 234 GlnGlnLeuLysSerLeuLeuGlyLeuGluHiPheThrHisGlyAlaAspIleLeuSer 253
DB 637 CAGCAGCTAAATAATATGTTGGGCATCAAGAGCTTTACAAAGAAACCAGATATAAATTTCT 696
QY 254 ValMetArgSerValPheThrGlnThrHisGlu---TrpArgTrpGluSerAlaValLeu 272
DB 697 GTTATGCGCTGCTGTTGGACTTCAGCTCATCAGGGTGGAAATGGCAAACTATTGTGATT 756
QY 273 GlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSerLysLysArgProArg 292
DB 757 GGCATAACTTTCCTTGCACTTCCTCTGCTGTCGCCAGTACATTGGAAAGAAAGATAGGAAG 816
QY 293 PhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuVal 312
DB 817 TTCTTCTGGGTGCGCAGCTATTGCTCTCTATTAACCTTCAGTTATTATTTGGCAACCCCTTTTGTG 876
QY 313 TyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysGlyLeu 332
DB 877 TTCACTACTCGTGTGACAGCAAGGTGTTTCAGATTGTTAAACCACATCAAAAAGGCATA 936
QY 333 AsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValLys 352
DB 937 AACCCATCATCTCCACAAAATTTATTTCACTGGTCCATTGTTGCCAAAAGGTTTCAAG 996
QY 353 ThrGlyIleValValGlyIleLeuSerLeuAlaGluGlyLeuAlaValGlyArgSerPhe 372
DB 997 ATCGGTGTCAATTCGCGCATGATCGGTTTAAACGGAAGCTGTGCAATTGGGAGGACGTTT 1056
QY 373 AlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMet 392
DB 1057 GCTGCTCTGAAGGACTATCAATTAGATGGAAACAAGGAGATGTGACACTTGGAACTATG 1116
QY 393 AsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAla 412
DB 1117 AACATAGCAGGGTCAATGACATCTTGTATATTGCAACAGGTCTTCTCTCGTTCGCA 1176
QY 413 ValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleLeuMetSerLeuAlaVal 432
DB 1177 GTCAATTTTCATGTGCTGCGTCCCAACACCAAGTGTCCAATATATATTATGTCAGCAGTTGTG 1236
QY 433 MetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSer 452
DB 1237 CTCTGACCTTGTGTGTCATCACCACTCTTCAAGTACACCAATGCGATCTCTCGGG 1296
QY 453 AlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPhe 472
DB 1297 TCGATAATCATTTCTGGGTGATCGGCTGTGACTACGAGCAGTAATCTCATTTGG 1356
QY 473 LysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValValPheGly 492
DB 1357 AAAGTTGACAAAATGGATTTCATTTCTTGCTGGGAGCATCTTTGGTGTGGTCTTCGCA 1416
QY 493 SerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuPhe 512
DB 1417 TCGTTGAGATAGGCTTTTAAATAGCTGTATGCAATTCATTTCTGCTAAAATTTCTTCTCAA 1476
QY 513 IleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsn 532
DB 1477 GTAACAAGGCCAAGGACTGTACTTCTGGGAACCTTCAGGCACCAATATACAGGAAC 1536
QY 533 ValGluHiSTyrGlnAsnAlaLysHisValProGlyMetLeuLeuLeuGluIleAspAla 552
DB 1537 ACTGATCAGTATCCAGAAGCAACACATATTTCTCTGGGGTGTATATAGTGGGTGATTCT 1596
QY 553 ProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGlu 572
DB 1597 GCTATCTACTTCTCCAACTAATGTCCGGGAGAGAACTCTTAGTGGTGTTCGGAG 1656
```

```
QY 573 GluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMet 592
DB 1657 GAAGAAGAAAGCAACACGAGGACAGTCTAAAATCAATTTCTTGTATCATAGAGATG 1716
QY 593 SerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLysLysIle 612
DB 1717 TCCCGGTTCATAGATATCGATCAAGTGGCATCCATCTCTAGAAGATCTATACAAGAAT 1776
QY 613 ThrGluArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLysLys 632
DB 1777 CTTAAGAAGAGACATTCAGCTCATCTCCTAGCAATCCTGGATCCATCGTCATGGAAG 1836
QY 633 LeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysLysTrpIleTyrLeuThrValGlu 652
DB 1837 CTCTTGTCTATCCAGCTCAACGAGCACATAGGAGTAAACAATATTTCTCCTACGGTTGCT 1896
QY 653 GluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerLysThrAsnPro 669
DB 1897 GACGCTGTGCTTCTGCACA-----AGAAAGTCGATCGTCAAGAACCA 1938

RESULT 13
ABQ82719
ID ABQ82719 standard; cDNA; 1971 BP.
XX ABQ82719;
XX 03-JAN-2003 (first entry)
DE Arabidopsis thaliana sulphate transporter Sultr1;3 encoding cDNA SEQ:1.
XX Arabidopsis thaliana; sulphate transporter; chromosome 1; plant;
KW sulphur-containing metabolite; glutathione; phytokeatin; gene; ss.
XX Arabidopsis thaliana.
XX Key Location/Qualifiers
FH 1. .1971
FT CDS
FT /*tag= a
FT /product= "sulphate transporter Sultr1;3"
XX
XX JP2002272472-A.
XX
XX 24-SEP-2002.
XX
XX 22-MAR-2001; 2001JP-00082891.
XX
XX 22-MAR-2001; 2001JP-00082891.
XX
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
XX WPI; 2002-718707/78.
XX P-PSDB; ABP53777.
XX
XX A gene encoding sulfate ion transporter Sultr1;3, a polynucleotide, an
oligonucleotide, a recombinant vector, a plant.
XX
XX Claim 2; Page 6-9; 12pp; Japanese.
XX
XX The present invention describes a gene which is present in the first
chromosome of Arabidopsis thaliana and encodes sulphate ion transporter
Sultr1;3 having the amino acid sequence given in ABP53777. Also
described: a polynucleotide purified from the genomic DNA, the mRNA, the
cDNA or their complementary sequences; an oligonucleotide hybridising
with the above gene or the above polynucleotide and consisting of a base
sequence of at least 10 base pairs; a recombinant vector carrying the
above polynucleotide; and a plant in which the above gene is introduced
or its descendant or their tissues. The plant can be used for the
preparation of sulphur-containing metabolites such as glutathione and
phytokeatin. The present sequence encodes Arabidopsis thaliana sulphur
transporter Sultr1;3, from the present invention
XX
XX Sequence 1971 BP; 504 A; 465 C; 427 G; 575 T; 0 U; 0 Other;
```

Alignment Scores:				
Pred. No.:	2.84e-176	Length:	1971	
Score:	1807.50	Matches:	335	
Percent Similarity:	72.76%	Conservative:	119	
Best Local Similarity:	53.69%	Mismatches:	169	
Query Match:	52.01%	Indels:	1	
DB:	6	Gaps:	1	
US-10-762-049-18 (1-680) x ABQ82719 (1-1971)				
Qy	37	AsnPheGluArgValHisGlnValGluValProProGlnProPhePheLysSerLeu	56	
Db	76	AATACACCGATGTCACCAAGTCGAAGTTCCTCTTAAGCAAAACCTTTTCAATGAGTTC	135	
Qy	57	LysTyrSerLeuLysGluThrPhePheProAspPheProLeuArgGlnPheLysAsnLys	76	
Db	136	ATGTACACTTTTAAAGAAACTTTCTCCAGCATGATCCTCTAAGGCATTTTCAAGGACAC	195	
Qy	77	ProAlaSerLysLysPheMetLeuGlyLeuGlnPhePheProLeuPheGluTtpAla	96	
Db	196	TCAAAATCCAAAGAGCTGCTCGGTATCCAGTCCGCTTTCCGGTTATCGAGTGGGGA	255	
Qy	97	ProLysTyrThrPheGlnPheLeuLysAlaAspLeuileAlaGlyIleThrIleAlaSer	116	
Db	256	AGAAATATAATCTTAAGTTGTTTCGCGCGATCTTATTGCGGGTTTAAACATAGCCAGT	315	
Qy	117	LeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGly	136	
Db	316	CTCTGCATTCCTCAGGATATGGATATGCAAGCTCGAGTCTTGACCCCTAAAGTAGGT	375	
Qy	137	LeuTyrSerSerPheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeu	156	
Db	376	CTATATTCAAGTTTGTCTCCGTGTTGTGATGCGATGTATGGGAAGCTCAAAAGGATATA	435	
Qy	157	AlaValGlyThrValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaVal	176	
Db	436	GCGATTGGACCGTTCAGTGGTTTCACTCTATTAGGTACTCTGCTCGTGCAGATC	495	
Qy	177	AspProAsnGluAspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAla	196	
Db	496	GACCCCAACCAAAACCTTAATGAATATCTCCGCTTAGCCTTACCTCCAGCTTCTTTGCC	555	
Qy	197	GlyValPheGlnAlaLeuGlyLeuPheArgLeuLeuIleValAspPheLeuSer	216	
Db	556	GGTGTCACTCAAGCAGCCTCGGATCTTCAGATTGGTGTTCGTGATTCGATTCTTGTC	615	
Qy	217	HisAlaThrIleLeuGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeu	236	
Db	616	CACGCGCGGTGGTAGTTCATCGGCGGAGCAGCCATCACCATTGCGCTGCACAGCTC	675	
Qy	237	LysSerIleLeuGlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArg	256	
Db	676	AAAGGCTTCCTAGGAATCAATAGTTTCAAAAGAAACCAGATATCATCGCGGTCTTTCT	735	
Qy	257	SerValPheThrGlnThrHisGlu---TtpArgTtpGluSerAlaValLeuGlyCysVal	275	
Db	736	TCCGTAAATCAGCTCAGCCCATCAGGATGAAATGGCAGACAATACTCATGTGTCATCG	795	
Qy	276	PheIlePhePheLeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePheTtp	295	
Db	796	TTCTTGATCTCTCTCTCATCTCCAGTTTATCGGGAAGAGAAACAAAGAACTGTTTGG	855	
Qy	296	ValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThr	315	
Db	856	ATTCAGCTATTGTCGTTAGTATCTGTATCATCTTCAACCTTCTTCGTCATACATAACC	915	
Qy	316	HisAlaGlyLysHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProPro	335	
Db	916	CGAGCCGACAGAAAGAGTTCAGATAGTGAACATCTTGACAAAGGCTCTGAACCCCTCT	975	
Qy	336	SerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIle	355	
Db	976	TCITTTGGCTTAATATATTATTCTCGGGCGATTTACCTTCTCAAGGGCTTCCGCATAGGCGTT	1035	

## RESULT 14

ADA70110

ID ADA70110 standard; DNA; 2000 BP.

XX

AC ADA70110;

XX

DT 20-NOV-2003 (first entry)

XX

Qy	356	ValValGlyIleIleSerLeuAlaGluGlyIleAlaValAlcGlyArgSerPheAlaMetTyr	375	
Db	1036	GTCTCAGGATGTTGCTTTCGCGAAGCTGTAGCATAGAGAACAATTTGCGCAATG	1095	
Qy	376	LysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValVal	395	
Db	1096	AAAGACTACCAATCGATGTAACAAGAGATGGTAGCATTAGAGCAATGAACGTAACT	1155	
Qy	396	GlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyr	415	
Db	1156	GGTTCAATGACCTTGTCTATGATCCACCGGTTCTTCTCAAGATCCGCGCAACTTT	1215	
Qy	416	AsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThr	435	
Db	1216	ATGCCCGGATGTCAAAACGCGAGTCTCCAACATCATCATGTGCTGCTCTTAAACG	1275	
Qy	436	LeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIle	455	
Db	1276	CTTCTCTTCTCACTCTCTCTTCAATATACACACCAACCAATTTTGGAGGATCATC	1335	
Qy	456	ValSerAlaMetLeuGlyLeuIleAspTyrGluAlaIleHisLeuPheLysValAsp	475	
Db	1336	ATCAACGCTGTGATTCCTTTGGTGTAGCTTAATGCTACCATTTTGATCTTCAAGATCGAT	1395	
Qy	476	LysPheAspPheValValCysMetSerAlaTyrIleGlyValValPheGlySerValGlu	495	
Db	1396	AAGCTCGATTGTTGCTGTATGGGGGCTTTTGGTGTCTATCTTGTATCCGTAGAG	1455	
Qy	496	IleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArg	515	
Db	1456	ATTGGGCTTCTAATAGCGGTGGGATATCTTTGCCAAGATATCTTTCAAGTTTACGAGA	1515	
Qy	516	ProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHis	535	
Db	1516	CCTAGGACAGCGATTCTTGGAAAGATACACAGGACTTCGGTTTACAGGAATATCAATCAG	1575	
Qy	536	TyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyr	555	
Db	1576	TATCTCTGAAGCGACTAGGATTCGGGAGTTTTCACAATTCGTGTGCTCCGCGATTAC	1635	
Qy	556	PheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTtpIleAspGluGluGluGlu	575	
Db	1636	TTCTCCACTCCCAATTTATGTTAGGAAAGGATTCAAAGATGGTTGACAGATGAAGAAGAG	1695	
Qy	576	ArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaVal	595	
Db	1696	ATGGTCGAAGCTGCAAGATTGCTAGGATCCAGTTCCTTATCATCGAAATGTCACTGTT	1755	
Qy	596	GlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArg	615	
Db	1756	ACGACATCGATACTAGTGTATTCAACGCTTAGAAGACTTGTATAGTCTCTCCAAAAA	1815	
Qy	616	ArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLys	635	
Db	1816	CGAGATTCAGTTGTTCTAGCGAATCCAGGACCCCGGTCAATAATAGCTACATGTT	1875	
Qy	636	SerLysPheGlnAsnHisLeuGlyLysTyrTtpIleTyrLeuThrValGluGluAlaVal	655	
Db	1876	TCCTCACTTTCGGGACTTGTAGGACACGACCAAAATCTTCTGACGGTGGCGCGGTG	1935	
Qy	656	GlyAlaCysAsn 659		
Db	1936	GATTCCTGCTCC 1947		

## RESULT 14

ADA70110

ID ADA70110 standard; DNA; 2000 BP.

XX

AC ADA70110;

XX

DT 20-NOV-2003 (first entry)

XX





Db	1486	GGCTTGTCAGTTGTCAGGGCGTTGCTGTACTGGCCAGGCGCTGCCACCTGCAAGCTCGGG	1544
Qy	523	AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVal	542
Db	1546	AACATAGCAGGCGAGGACCTTCGCGGACGTGAAGCAATACCCCAAGCGNAGAGCATC	1605
Qy	543	ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu	562
Db	1606	CCTGGCATCCCTGCTTTCGACCTCGGCTCTCCCATCTACTTCGTCAATGGGGTTACCTG	1665
Qy	563	ArgGluArgIleThrArgTTrIleAspCpIuGluGluArgIleLysAlaThrGlyGlu	582
Db	1666	CGAAGAGATTTCAGATGGTGGAGATGAGGATACTGTGCAAGAGCGTCGGGCAC	1725
Qy	583	ThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly	602
Db	1726	--GATCTGCAATACTTGGTTCTTGATCTTGGTGGTGCATCTTCTGTCGACACTCTGA	1782
Qy	603	IleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeu	622
Db	1783	GTTGGGATGCTACTAGAAGTACACAGAGCGCTCGAACGAGAGGATCACCATAGTCTCTG	1842
Qy	623	ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe--GlnAsnHis	641
Db	1843	ACGAATCCGAGGCTAGAGGTAAACAGAGAAGCTGGTGTCTGTGATACGTACGGGACATC	1902
Qy	642	LeuGlyLysLysTrrIleTyrLeuThrValGluAlaValGlyAlaCysAsnPheAsn	661
Db	1903	TTAGGGGATGAGTGGTCTTCTTGACCGTCAAGGACGCCATCACGGGCTTCGATAGCGG	1962
Qy	662	LeuArgAlaSerLysThrAsnProLysLysAspGlu	673
Db	1963	CTGCAGATATCCAGA--AATAAGGAGAGACGAA	1995
RESULT 15			
Id	AAZ50490 standard; cDNA; 2311 BP.		
XX	AAZ50490;		
XX	23-MAY-2000 (first entry)		
DE	Wheat sulphate permease-1 cDNA clone.		
KW	Sulphate Permease; sulphate assimilation protein; wheat; probe; mapping;		
KW	marker; plant breeding; chimeric gene; transgenic plant; antibody;		
OS	Triticum aestivum.		
Key	Location/Qualifiers		
FT	100..2070		
FT	/*tag= a		
FT	/product= "Wheat sulphate permease-1"		
FT	/note= "Derived from clone wiki.pk0028.e1"		
XX	WO200004154-A2.		
XX	27-JAN-2000.		
XX	13-JUL-1999; 99WO-US015810.		
XX	14-JUL-1998; 98US-0092833P.		
PA	(DUPO ) DU FONT DE NEMOURS & CO E. I.		
XX	Allen SM, Falco SC, Thorpe CJ;		
XX	WPI; 2000-195025/17.		
DR	P-PSDB; AAY44944.		
XX	Nucleic acid fragments encoding sulfate assimilation proteins in plants		
PT	and seeds useful as probes for isolating cDNAs and genes encoding		

homologous proteins, in producing transgenic plants.

Claim 3; Page 55-56; 79pp; English.

The present sequence is a cDNA clone encoding wheat sulphate permease, a sulphate assimilation protein. This sequence is obtained from wkl1 library, clone wkl1.pk0029.e1, derived from wheat seedlings, 1 hour after treatment with fungicide. This sequence is used as a probe to isolate other plant sulphate assimilation proteins, for genetic and physical mapping of related genes and as markers of traits linked to the gene. This is useful for plant breeding and to construct chimeric genes, used to create transgenic plants with altered levels of sulphate permease. The sulphate permease peptides are useful for producing antibodies, that are used to screen and isolate cDNA clones

Sequence 2311 BP; 670 A; 493 C; 537 G; 611 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	2,73e-175	Length:	2311
Score:	1799.00	Matches:	336
Percent Similarity:	72.30%	Conservative:	126
Best Local Similarity:	52.58%	Mismatches:	172
Query Match:	51.77%	Indels:	5
DB:	3	Gaps:	2

US-10-762-049-18 (1-680) x AAZ50490 (1-2311)

Qy	41	ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu	60
Db	187	GTGTACAAAGTGGGCTATCCCCCTCCGAAGACTTGGCCACAGAGTTTACAGAAACATTTG	246
Qy	61	LysGluThrPhePheProAspProLeuAtgGlnPheLysAsnLysProAlaSerLys	80
Db	247	AGGGAGACTTTCCTCCACGACACCCGCTGGCTCAGTATAAGGGCCCAATCCGACCGAGG	306
Qy	81	LysPheMetLeuGlyLeuGlnPhePheProPheGluThrPalaProLysTyrThr	100
Db	307	AGGTTTCATGATGGGCTGGAGTTCTTGTTTCCATATTTGGGTGGGTATGAGTACAGT	366
Qy	101	PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro	120
Db	367	CTCACAAAGTTCAAAGGGCATCTGATGCGGATGACCATCGCAAGTCTCTGTATTCCT	426
Qy	121	GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProPheLeuGlyLeuTyrSerSer	140
Db	427	CAGGACATTGGCTATTTCGAAGCTTGTATCTGGATCCGACAGTATCGGCTTTACTCCAGC	486
Qy	141	PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr	160
Db	487	TTCAATTCCTCCATTGATCTATGCTGCAATGGGTAGCTCAAGGGATATAGCGATTGGTCCA	546
Qy	161	ValAlaValGlySerLeuMetGlySerMetLeuSerAsnAlaValAlaProAsnGlu	180
Db	547	GTTCGTGTGGTTCTCTTTGTAGAGTTCACTTTACAAGCTGAGGTGACCATGTCAAA	606
Qy	181	AspProLysLeuTyrLeuHisIleuAlaPheThrAlaThrLeuPheAlaGlyValPheGln	200
Db	607	AACAAGGAGGAATACATCGCCTCGCTTTCACGGCAACCTTCTTCGCTGGTATCATCTCAA	666
Qy	201	AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLysSerHisAlaThrIle	220
Db	667	GCAGCCTTAGGATTTCTTAAGGTTAGGATTCCTTATAGAGTTCTTGTTCGATGCTGGATT	726
Qy	221	IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu	240
Db	727	GTTCGGATTATGCGGGGAGCTGCATTAATTATTCCTTCGATGAGTGAATACGTGTTG	786
Qy	241	GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr	260
Db	787	GGCATCGCAAACTTTACAGAGAAACCGACATAGTTTCTGTCATGGAATCTGCTGGAGA	846
Qy	261	GlnThrHisGlu---TrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePhe	279



847	Db	TCAGTTTCATCATCGGGTGGAACTGGCAGACAAATGTTGTGATGTTGGCGGTATCTTTCTCTGGTTTTC	906
280	Qy	LeuLeuSerThrArgTyrPheSerLysArgProArgPhePheTyrPValSerAlaMet	299
907	Db	CTTCTGTTGCGAAGTACATCGGAAGAGAGAAAGGAAGCTTTTCTGGGTGCCAGCTATT	966
300	Qy	AlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLys	319
967	Db	GCTCCTATAATTTCTCAGTCGATTCTAGCAACATTTTTTGTATACATTTACTCGTGGCGACAAG	1026
320	Qy	HisGlyValGluValIleGlyGluLeuLysGlyLeuAsnProProSerLeuThrAsn	339
1027	Db	CAAGGAGTTCAGATAGTGAAGACATTCGAACAGGGAATCAACCCATCATCAGTACACAAG	1086
340	Qy	LeuValPheValSerProTyrMetThrThraValLysThrGlyIleValValGlyIle	359
1087	Db	ATTATTATTCACGGCCCATTTGTTGCAAAAGGTTTCAAGATCGGTGTTGTTTGGCGGATA	1146
360	Qy	IleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsn	379
1147	Db	GTTGGTTTGACAGAAGCTGTAGCTATTGGAAGGACATTTGCTGCTATGCAAGGACTACCAG	1206
380	Qy	IleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThr	399
1207	Db	TTAGATGGAACAAGSAGATGTTAGTACATTTGGAAACCATGAACATGAGTGGCTCAATGACA	1266
400	Qy	SerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCys	419
1267	Db	TCTTGCTATGTCACACACAGTTCTTTCTACGTTTCGGCAGTTTAACCTTCATGCGCTGGCTGC	1326
420	Qy	LysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeu	439
1327	Db	AAGACTCCTGTATCCATGTTGTTATGTCAGTAGTGGTTCTCTTACCTTGGTGGTCAATC	1386
440	Qy	ThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMet	459
1387	Db	ACACCGCTATTCAAATATACACCGAATGCAATCCCTAGGTCGATCATTTATTTCTGCGGTG	1446
460	Qy	LeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPhe	479
1447	Db	ATCGGCTTTGGACTACGAAGCAGCAATTCATCTGGAAGTGTGCAAAATTTGGAATTC	1506
480	Qy	ValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVal	499
1507	Db	ATTGCTTGCATGGGAGCTTTTTCGGTGTGTTTGTATCCGTTGAGATTGGCTCTTG	1566
500	Qy	IleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPhe	519
1567	Db	ATTGCTGTAGCAATCTCATTTGGCCAAATACTTCTTCAAGTAAACAGGCCCAAGCAGACC	1626
520	Qy	ValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAla	539
1627	Db	CTACTTGGAAACCTTCCCGGCACCCTATATACCGGAACATCAGCCAGTATCCAGAAGCA	1686
540	Qy	LysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAla	559
1687	Db	AAACTTACTCTGGGGTGCATGTGAGGGTGTGATTCTGCTATTATTTTCCACTCT	1746
560	Qy	SerTyrLeuArgGluArgIleThrArgTyrPileAspGluGluGluArgIleLysAla	579
1747	Db	AAATTACGCGAGAAAGAAATTTCTTAGTGCGCTGACACACGCAAGAACAGAGCTAAAGCA	1806
580	Qy	ThrGlyLuthrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAsp	599
1807	Db	GTGGGATTGCTAAAATCAGTTTCTGATTGTGGAATGTGCGCGGTATCGACATCGAT	1866
600	Qy	ThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGln	619
1867	Db	ACAAGCGCATACATGCTCTTGAAGATCTATACAGGAATCTTCAGAAAAAAGATATGCGAG	1926
620	Qy	LeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGln	639
1927	Db	CTCAATCTGTGCAATCTCGTTCGTCTATAGAAAAATCGCAAGCGTTCGAGCTCACC	1986

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 31, 2005, 11:52:46 ; Search time 331 Seconds  
(without alignments)  
3361.535 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLLRNVTI.....NLRASKTNPKDETEGNNV 680

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10762049/runat\_30082005\_121246\_6006/app\_query.fasta\_1.839  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THRE SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10762049 @CGN 1.1.69 @runat\_30082005\_121246\_6006 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued Patents NA.\*
- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3475	100.0	2449	4	US-09-720-317A-17
2	2164	62.3	1981	4	US-09-720-317A-3
3	1901	54.7	2067	4	US-09-720-317A-15
4	1844.5	53.1	2279	4	US-09-720-317A-1
5	1799	51.8	2311	4	US-09-720-317A-19
6	963.5	27.7	2022	4	US-09-720-317A-21
7	731.5	21.1	1780	4	US-09-720-317A-7
8	687.5	19.8	1767	4	US-09-252-991A-7241
9	676.5	19.5	1767	4	US-09-252-991A-7184
10	675.5	19.4	1240	4	US-09-720-317A-5
11	651	18.7	2441	4	US-09-785-381-4
12	650	18.7	4113	4	US-09-785-381-2

13	643.5	18.5	1773	4	US-09-902-540-7746	Sequence 7746, Ap
14	643.5	18.5	4854	4	US-09-902-540-768	Sequence 768, App
15	641.5	18.5	2832	4	US-09-949-016-4516	Sequence 4516, Ap
C 16	625	18.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 17	619	17.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
18	612	17.6	1818	4	US-09-602-787A-615	Sequence 615, App
19	586.5	16.9	2858	4	US-09-949-016-5349	Sequence 5349, Ap
20	586.5	16.9	2882	1	US-08-424-567-1	Sequence 1, Appli
21	586.5	16.9	2882	2	US-08-711-928-1	Sequence 1, Appli
22	586.5	16.9	2882	3	US-09-184-937-1	Sequence 1, Appli
23	578	16.6	2682	4	US-09-614-221A-222	Sequence 222, App
24	577	16.6	4927	4	US-09-949-016-5627	Sequence 5627, Ap
C 25	571.5	16.4	31147	4	US-09-596-002-25	Sequence 25, Appl
C 26	567	16.3	2487	4	US-09-248-796A-6643	Sequence 6643, Ap
27	548.5	15.8	1992	4	US-09-875-811-5	Sequence 5, Appli
28	548	15.8	1971	4	US-09-875-811-9	Sequence 9, Appli
29	548	15.8	2420	4	US-09-875-811-13	Sequence 13, Appli
30	545.5	15.7	2040	4	US-09-875-811-1	Sequence 1, Appli
31	499	14.1	1851	4	US-09-543-681A-249	Sequence 249, App
32	485.5	14.0	1818	4	US-09-875-811-7	Sequence 7, Appli
33	485	14.0	1797	4	US-09-875-811-11	Sequence 11, Appli
34	482.5	13.9	1866	4	US-09-875-811-3	Sequence 3, Appli
35	464	13.4	8774	4	US-09-949-016-16258	Sequence 16258, A
36	451	13.0	2913	4	US-09-795-927-6	Sequence 6, Appli
37	451	13.0	3749	4	US-09-795-927-8	Sequence 8, Appli
38	449	12.9	1608	4	US-09-252-991A-5234	Sequence 5234, Ap
39	449	12.9	1911	4	US-09-252-991A-5199	Sequence 5199, Ap
40	426.5	12.3	1419	4	US-09-902-540-6127	Sequence 6127, Ap
41	426.5	12.3	2070	4	US-09-902-540-286	Sequence 286, App
C 42	402.5	11.6	1230230	4	US-09-438-185A-1	Sequence 1, Appli
C 43	389.5	11.2	1230025	4	US-09-198-452A-1	Sequence 1, Appli
C 44	372	10.7	1116	4	US-09-252-991A-7460	Sequence 7460, Ap
C 45	369	10.6	1701	4	US-09-543-681A-372	Sequence 372, App

ALIGNMENTS

RESULT 1  
US-09-720-317A-17  
; Sequence 17, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 17  
; LENGTH: 2449  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-720-317A-17

Alignment Scores:  
Pred. No.: 0  
Score: 3475.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 4  
Length: 2449  
Matches: 680  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-10-762-049-18 (1-680) x US-09-720-317A-17 (1-2449)

Qy 1 HisGluLeuAlaArgThrLeuSerTyrIleThrHisCysLeuLeuArgAsnThrIle 20  
Db 2 CACGACTAGCTCCACANTTAGTTATATACACATATTTCCTTGCTTAGAATATTATT 61



; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 3  
; LENGTH: 1981  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-720-317A-3

## Alignment Scores:

Pred. No.: 2,17e-258 Length: 1981  
Score: 2164.00 Matches: 403  
Percent Similarity: 86.09% Conservative: 86  
Best Local Similarity: 70.95% Mismatches: 79  
Query Match: 62.27% Indels: 0  
DB: 4 Gaps: 0

US-10-762-049-18 (1-680) x US-09-720-317A-3 (1-1981)

```
QY 105 LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer 124
DB 8 GAGTCGACCTGATCGCGCGCATCACCATCGCCAGCCTCGCCATCCGCGCAGGGGCATCAGC 67

QY 125 TyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIleProPro 144
DB 68 TAGCCCAAGCTCGCCCAACCTGCCGCCCGCTGCTCGGACTCTACTCGAGCTTCGTGCCGCCG 127

QY 145 LeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGly 164
DB 128 CTGGTGATCGCTGATGGGAGCTCCCAAGACCTGGCGGTGGGACCGTGGCGGTGGCG 187

QY 165 SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeu 184
DB 188 TCGCTGCTCATAGCTCCATGCTCGGCAGCGAGGTGTCGCCGACGGAGAACCCCGTGCTC 247

QY 185 TyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGly 204
DB 248 TACCTGCACCTCGCCTTACCGCCACCTTCTTCCCGCGGCTTCTCGAGGCTCGCTCGGC 307

QY 205 LeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMet 224
DB 308 CTCCTCAGGTTGGGCTTCATCGTGACCTGCTGTCGCACGCGACGATCGTGGGTTTCATG 367

QY 225 GlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis 244
DB 368 GCCGCGCGCGCAGCGGTGTGCTCGCAGCAGCTGAAGGCGCATGCTGGGCGCTCGTCCAC 427

QY 245 PheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGlu 264
DB 428 TTACACCACTCCACCGACGCTGCTCGTATGGAATCCGTTCTCAGCCACACACACAG 487

QY 265 TrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArg 284
DB 488 TGGCGGTGGGAGACGCTGCTCGGCTGGCGCTTCTCTTCTCTCTCTCTCTCTCTCTCTCT 547

QY 285 TyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSer 304
DB 548 TTCAATCAGCAAGAGCGCTCCCAAGCTGTTTGGATCTCCGCGCGCGCGCTTGAAGCTCC 607

QY 305 ValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluVal 324
DB 608 GTCGTGTCTGGGAGCTTGTGTGTACTCAGCAGCAGCTGANAACACACGCGCATCGAAGTG 667

QY 325 IleGlyLeuLeuLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSer 344
DB 668 ATCGGTTTACCTGAAGAAGGCTGAATCCACCGTCGGTGACAGCCTGCAATTCTCACCG 727

QY 345 ProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGlu 364
DB 728 CCTTACATGATGCTCGCGGCTCAAGACTGGGATCATCCCGCGTCATTCGCTCGCGCAA 787

QY 365 GlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLys 384
DB 788 GGAATCCCGTGGGAGGAGCTTCGCCATGTTCAAGAACTTACACATGACGGACACCAAG 847
```

```
QY 385 GluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThr 404
DB 848 GAGATGATCGCGATCGGACGATGAACGCTCGCTCGGCTCGCTACGCTGCTACTCGTACC 907

QY 405 ThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSer 424
DB 908 ACGGGGCGCTTCTCGCGCTCCGCGGTAACCTACACGCCGGTGCAGGCGCATGCTCG 967

QY 425 AsnIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHis 444
DB 968 AACGTGTCATGTCGCTGGCGGTGATGTCACGCTGCTTCTGACGCCGCTGTTCCAC 1027

QY 445 TyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAsp 464
DB 1028 TACACGCCGCTGCTGCTGTCGCGCATCATGCTCTCCGCGATGCTGGGCTGGTGCAC 1087

QY 465 TyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSer 484
DB 1088 TTCCGGGCGCGCTGCACCTGTGCGCGCTGCAGAGTCAGCTTCTGCGTCTGCGCGCGC 1147

QY 485 AlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIle 504
DB 1148 GCCTACCTCGCGCTGCTTTCGGCAGCGTCGAGGTGCGCTGCTGCTGCGCGTCCGCTC 1207

QY 505 SerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIle 524
DB 1208 TCCTGCTCGCGCTGCTGCTGCTGCTGCGCGCGCCAGGACCGTGTCTCGCAACATC 1267

QY 525 ProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGly 544
DB 1268 CCCGCCACCATGCTGACCGAGGATGGACCATGACCCCGCGCGCAGAGCGTGCCTCGC 1327

QY 545 MetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGlu 564
DB 1328 GTGCTGCTGCGCTGCGCTGCGCGCTGCTTCTTCCCAACGCCAGCTACCTGCGGAG 1387

QY 565 ArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSer 584
DB 1388 AGGATCTCGCGGTGGATCGACGACGAGGAGGAGCGCACCAAGAGCCAGGCGGAGTGGC 1447

QY 585 LeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSer 604
DB 1448 GTCCGTCAGTGTCTCTCGACATGGGTGCCATCGGTAGCATCGACAGCGGCGACGAG 1507

QY 605 MetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsn 624
DB 1508 ATGCTGACGAGCTCAACAGTCTTGGACAGGAGGGAATGCGATCGTGTGCGCGAAC 1567

QY 625 ProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLys 644
DB 1568 CCGCGCAGCGAGATCATGAAGAAGCTGGAACAGCTCCAAAGGTGTGGAGCAGATCGGCG 1627

QY 645 LysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAla 664
DB 1628 GAGTGGGTGTCCCGAGCGTGGCGGAGGCGGTGCTGTCGCGCATCTGCTGCTGCTGCTG 1687

QY 665 SerLysThrAsnProLysLysAsp 672
DB 1688 CACAAGCCGGGAATGGCCAAAGGAC 1711
```

## RESULT 3

US-09-720-317A-15  
; Sequence 15, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21

```
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 2067
; TYPE: DNA
; ORGANISM: Glycine max
US-09-720-317A-15

Alignment Scores:
Pred. No.: 1,31e-225 Length: 2067
Score: 1901.00 Matches: 359
Percent Similarity: 76.48% Conservative: 106
Best Local Similarity: 59.05% Mismatches: 143
Query Match: 54.71% Indels: 0
DB: 4 Gaps: 0

US-10-762-049-18 (1-680) x US-09-720-317A-15 (1-2067)

QY 48 ProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGluThrPhePheProAsp 67
DB 8 CCACACAGACCACTCCACAACTTAGCCACAGAGTCTCCGAAATCTTCTCCCGAGAT 67
QY 68 AspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPheMetLeuGlyLeuGln 87
DB 68 GACCCCTCCACCGTTTCAAGAACCAAACTCGCTTTAAAAAGTTCTCTCGCACTTCAG 127
QY 88 PhePhePheProLeuPheGluThrAlaProLysTyrThrPheGlnPheLeuLysAlaAsp 107
DB 128 TATCTCTTCCCAATTTTCGACTCGGCCCCCAAACTACAATCTTACCCTTCTCGCTCTGAC 187
QY 108 LeuAlaGlyLeuThrLeuAlaSerLeuAlaIleProGlnGlyLeuSerTyrAlaLys 127
DB 188 CTCATCTCGGCTCAACATTCAGCAGCTGCCATTCCTCAGGGAATCAGTTATGCCAAG 247
QY 128 LeuAlaAsnLeuProProLeuLeuGlyLeuTyrSerPheIleProLeuIleTyr 147
DB 248 CTTCGCAACTTGCACCTATTCTGGATTATATTCGAGTTTGTTCCTCCCATGATATAC 307
QY 148 AlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeu 167
DB 308 TCCTCTCTTGGAGATTCTAGACATCTGGTGTGGACCTGTTTCCATTCGCTCTTGGTC 367
QY 168 MetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHis 187
DB 368 ATGGGATCAATGTAGTATGATAAAATCTTACACTCAAGAACTATTCTCTATCTGGGA 427
QY 188 LeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArg 207
DB 428 TTGGCTTTACCGCCACTTCTTTGCTGGTGTATTCCAAGCTTCTCTGGGTATATTAAAG 487
QY 208 LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleLeuGlyPheMetGlyAla 227
DB 488 CTAGGCTTCGTAATGATTCTTCGAAAGGCAACGCTGTTGGATTCAACGCGGTGCT 547
QY 228 AlaThrValValCysLeuGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHis 247
DB 548 GCCATTATGTGTCACTGCGACGAGCTGAAAGGTTTACTTGGAAATAGTGACCTTACAGC 607
QY 248 GlyAlaAspIleLeuSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrp 267
DB 608 AAGATGCAAAATAATTCAGTAACGATCTCTGTTTCAAGCAAAAGACACGAGTGTCTATGG 667
QY 268 GluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSer 287
DB 668 CAAACCATTTCTTTGGGATTCGGCTTCTCTGCTGATCTTCTGCTGACAAAGGACATTAGT 727
QY 288 LysLysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeu 307
DB 728 TTGAGGAACCAAACTATTCTGGGTTTTCAGCAGCTGCCCATTCAGTATCTCTG 787
QY 308 GlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGlu 327
```

RESULT 4

US-09-720-317A-1

```
DB 788 TCAACCACTTTAGTCTTTCTTGAGAAATAAGACTCATCAAAATTCAGTTATTGGGCAC 847
QY 328 LeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMet 347
DB 848 TTACCAAAAGGGAGTTAAATCCACCATCAGCAAAATGTTTATCTCAATCGTCTTACTTG 907
QY 348 ThrThrAlaValLysThrGlyIleValValGlyIleLeuSerLeuAlaGluGlyIleAla 367
DB 908 GGTCTTGCTATCAAAACTGGCATCATCAGAGGATCTTATCTCTCACTGAAGGAATTGCA 967
QY 368 ValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIle 387
DB 968 GTAGGAGAAACATTTGCTTCACTTAAGAACTACCAGGTGATGGAACAAGAAATGATG 1027
QY 388 AlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrGlyPro 407
DB 1028 GCCATTGGTCTAATGAACATAGCTGGCTGCTGTTCTTCATGTTATGTTTACAACGGGATCC 1087
QY 408 PheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIle 427
DB 1088 TTTTCTCGATCGCTGTGTTAACTATATGCTGGAGCACAGACAGTTTCAAAATATAATC 1147
QY 428 MetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrPro 447
DB 1148 ATGGCTGCAGCTGTCTAGTGACACTTCTGTCTTCATGCTCTCTTCTTACTATACACCA 1207
QY 448 LeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAla 467
DB 1208 AATGTTGCTTAGCGCCCATTTATCATCTGCTGTGATTTGGTCTAATAGATTATCAATCT 1267
QY 468 AlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIle 487
DB 1268 GCATATAAATTTGGAAGGTTGACAACTTGATTCTTGGCCTGTTTGGCTCTTTT 1327
QY 488 GlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeu 507
DB 1328 GGGGCTTCTGTTTCATTTCACTGCTGCTTTAGTCTCTGTTAGCGGTTTATCATATCAGTCTC 1387
QY 508 ArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSer 527
DB 1388 AAGATCTGCTTCATGTCCTCGACCAAACTTTGGGTTTTGGGAATATACAGGAACA 1447
QY 528 ValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIle 547
DB 1448 CAAATATCCACACATAAACCAATACAAAAAGCTTTAAGAGATTCCTTTCATTCTCAT 1507
QY 548 LeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThr 567
DB 1508 TTGGCTGTTGAGTCTCAATCTATTTTGTAACTCAACTTATCTTCAAGAAAGGATAC 1567
QY 568 ArgTTPLeuAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyr 587
DB 1568 AGATGGGTTTCGAAAGAGGAAGAGCATATAAAGCTATAATGGAGCTCCATTTGAAGTGC 1627
QY 588 ValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 607
DB 1628 ATAATTTTACATCATCAGCTGTCACAGCCACAGACACAAAGTGGGCTTGACACTTTATGT 1687
QY 608 GluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProValSer 627
DB 1688 GAACTTAGAAAGATGCTGGAGGAAGAGATCAGTTGAGTTTGTCTGCGCAAAATCCTGTGGA 1747
QY 628 GluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysTyrIle 647
DB 1748 AATGTGATGGAATAATTCATTAAGTCAAACTTTTGGATTCTTTTGGATTAAGAGGATC 1807
QY 648 TyrLeuThrValGluGluAlaVal 655
DB 1808 TATCTCACAGTGGGAGAGCTGTG 1831
```

```

; Sequence 1, Application US/09720317A
; Patent No. 696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2279
; TYPE: DNA
; ORGANISM: Zea mays
US-09-720-317A-1

Alignment Scores:
Pred. No.:      1,69e-218      Length:      2279
Score:          1844.50        Matches:     355
Percent Similarity: 74.25%      Conservative: 115
Best Local Similarity: 56.08%    Mismatches:  158
Query Match:     53.08%        Indels:       5
DB:              Gaps:         3

US-10-762-049-18 (1-680) x US-09-720-317A-1 (1-2279)

QY      41  ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
      189  GTGCACAGGTGGCGCGCCGACCGCGCGGAGCACCGCGAGCAAGATGAAGGTGAGGGTG 248
QY      61  LysGluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
      249  AAGGAGACCTCTCTCCCGACACCCGTTCCGGGGCGTTCAAGGGGCGACCGCGGGGACG 308
QY      81  LysPheMetLeuGlyLeuGlnPhePheProPheGluTrpAlaProLysTyrThr 100
      309  CAGTGGCTCATGGCGGTGAGGTACCTCTTCCCATCTGGAGTGGGTGGCGAGCTACTCC 368
QY     101  PheGlnPheLeuLysAlaAspLeuLeAlaGlyLeThrIleAlaSerLeuAlaIlePro 120
      369  TTGTGCTCTTCAAGTCCGACCTCGTCGCGGCGCTCACCATTGCAGCTCGCCATTCTCT 428
QY     121  GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProLysLeuGlyLeuTyrSerSer 140
      429  CAGGGCATTTAGCTAGCGGAAGCTGCGAAGCTTGCCATCTCCCATTAATTCGGGCTGT 488
QY     141  PheIleProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
      489  TTGCTGCCCGCATGGGTGTACCGGTGCTGGGGAGCTCCCGTGACCTGGCGGTGGGCCCG 548
QY     161  ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
      549  GTGTGATCTCTGCTGATCATGATGAGGTCCATGCTCGGCGGAGCGGTGAGCCCCACTGCG 608
QY     181  AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
      609  GAGCGACGCTTTCCTCGAGCTGGCCCTTCACTCCACCCCTGTGTGGCGGGGCTGGTGCAG 668
QY     201  AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
      669  GCCTCCTGGGCATCTCAGGCTCGGCTTGTGTATCGACTTTCCTGTCAAGGGGAGCGCTG 728
QY     221  IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
      729  GTGGGTTTCATGCCGCGCCGCCATCATCTGTGGCGCTGCAGCAACTCAAGGGGCTGCTG 788
QY     241  GlyLeuGluHisPheThrHisGlyAlaAspIleLeSerValMetArgSerValPheThr 260
      789  GGCATCGTCCACTTCACCAACGAGATGGCATCGTCCAGTCATGGCTCGCTCTTCCAC 848

```







1687	AAACTTACTCTCGGGTGGTGAATTTGTGAGGGTTGATTCTGCTATTTATTTTCCAACTCT	1746
560	SerTyrLeuArgGluAraGlleThrArgTrrpileAspGluGluGluAraGlleLysAla	579
1747	AAATTACGTCGAGAAGAANAATCTTAGTGGTGTCACACGAGNAGACAGAGCTAAACCA	1806
580	ThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerLaValGlyAsnIleAsp	599
1807	GTGGGATTCCTAAAAATCAGTTTCTCATTTGTGAAAATGTCGCGGTTCATCGACATCGAT	1866
600	ThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGln	619
1867	ACAAAGCGGCATACATGCTCTTTGAAGATCTATACAGAATACTTCAGAAAAAAGATATATGCAG	1926
620	LeuValLeuValIleAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGln	639
1927	CTCATTTCTGTCGAATCTCTGGTTCCGTTCATAGAAAACTGCAAGCGTCGAAGCTCACC	1986
640	AsnHisLeuGlyLysLysTrrpileTyrLeuThrValGluGluAlaValGlyAlaCysAsn	659
1987	GAGCACATTGGAAACGACCAATATATTCTCGCGGTCTCTGACGCTGTGCGATCTCTGTACG	2046
660	PheAsnLeuArgAlaSerLysThrAsnPro-LysLysAspGluThrGluGlyTrrp	677
2047	-----ACGAAGTCGATCGACGAACCGTCAGCGAAATGATGTTCGAGCAAAATGG	2092

## RESULT 6

```

US-09-720-317A-21
; Sequence 21, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-720-317A-21

```

Alignment Scores:		
Pred. No.:	6,819-109	Length:
Score:	963.50	Matches:
Percent Similarity:	55.33%	Conservative:
Best Local Similarity:	34.62%	Mismatches:
Query Match:	2.77%	Indels:
DB:	4	Gaps:
		2022
		206
		129
		243
		17
		7

US-10-762-049-18 (1-680) x US-09-720-317A-21 (1-2022)

Qy	91	ProIlePheGluTrpAlaProLysTyrThrPhe--GlnPheLeuLysAlaAspLeuIle	109
Db	66	CCGTGCTTGGCGTGGATGGGAGCTACAGATGGGAAGGAGGACTTCCAGGCCGACCTCGCC	125
Qy	110	AlaGlyIleThrIlealaserylLeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAla	129
Db	126	GC CGGCGATCAGCTGTGGCGTCATGCTTGTGCTCAGGCAATGTCTATATGCAAAAGCTGGCT	185
Qy	130	AsnLeuProProlIleLeuGlyLeuTyrSerSerPheIleProProLeuIleTyrAlaMet	149
Db	186	GGCGTTCACCCAAATTATGGGCTCTACACAGCGCTTGTGCCACTATTTGTCTCAGCGATT	245
Qy	150	MetGlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMetGly	169
Db	246	TTTGGTCTCTCAGCAAAATTACAGTAGTCCAGTGGCACTGTCTCTCTGTAGTAGTCC	305

Qy	170	SerMetLeuSerAsnAlaValAlaProAsnGluAspProLysLeuTyrLeuHisLeuAla	199
Db	306	AAATGTTCTTGGGGGTATAGTT-----AATTCATCTAGTGAAGCTGTACACGGAATTAGCC	359
Qy	190	PheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuGly	209
Db	360	ATATTATGGCATTCATGGTGGAAATCTGGAATGCTGTGATGCATTCGTAAGACTTGGC	419
Qy	210	LeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThr	229
Db	420	TGGCTATTTCGTTTCATAGCCATTCGTAAATATCGGATTCACACAGCTTCGGCCATC	479
Qy	230	ValValCysLeuGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAla	249
Db	480	GTAATGGTTGTGCCAAATCAAGTATTTCTGGGTACAGT---GTTACAAGAAGTAGC	536
Qy	250	AspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrpGluSer	269
Db	537	AAAATTATACCACTTATAGAGATATAATTGCTGGGAATAGATCAGTTCTCTCGCCCTCCA	596
Qy	270	AlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSerLysLys	289
Db	597	TTTGTAATGGGATCAGCGTTTCTTGTTATTCTTCTAATAATGAAAAAGCTAGGGAAAAACA	656
Qy	290	ArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGlySer	309
Db	657	AATAAAAAATTACGTTTCTCGAGAGCTTCGTGCCACTAACAGCTGTGTTCTTTGGAACA	716
Qy	310	LeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyLeuLys	329
Db	717	TTGTTTGTGAAAAATTTTTCGTCCTCAACT-----GCCATATCAGTGTGAGTGAATATCCG	770
Qy	330	LysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerPro-----	345
Db	771	CAAGGCCCTT-----CCAGT-----TTCTCCATTCTCGAGGATTTGAA	809
Qy	346	TyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGly	365
Db	810	CATCTGATGTCCTAATGCCAACTGCAATACTATACACTGGTGTCTATTTTGGAGTCT	869
Qy	366	IleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGlu	385
Db	870	GTTGGGATTCCTAAAGCGTTAGCTGCGAAGAAATGGTTATGAGTTGGACTCAAAACAAGAG	929
Qy	386	MetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThr	405
Db	930	TTATTGGCCCTTGGCTTATCAAAATATATGCGGTTCAATCTTCTCGCATATCTCGTACA	989
Qy	406	GlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsn	425
Db	990	GGCTCCTTTCTAGTCTGCTGTGATCATGAAGCGGGCAAGAACTGGATTATCAGGA	1049
Qy	426	IleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyr	445
Db	1050	ATCATATGGGCATAATAATTTCAGTGCTCTCTGTTTATGACACCAATATTACTGAT	1109
Qy	446	ThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyr	465
Db	1110	ATACCTCAGTGTGCATTTGGCTGTCATTTGATTTCTGTCTACCTGGCTGGTAGATTAT	1169
Qy	466	GluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAla	485
Db	1170	GAAAGGCCCATCTCTCTGGGGGTATTGATPAGAGAGATTTCTTTCTGGGGCGATGACA	1229
Qy	486	TyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSer	505
Db	1230	TTTTACTACACCTTAACCTTTTGGCATTTAGATTGGTGTCTCTGTTGGGTTCGGGTTTCG	1289
Qy	506	ValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIlePro	525
Db	1290	CTGGCATTTTGTGATCATCAATCTGCAATCCGCATATAGCTGTGTTTGGGCGGTTCCT	1349

```
QY 526 AsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMet 545
Db 1350 GGCACCACTGTTACAGGAATACATTGCAGTCCCTGAGGCTTATACATACACGGGATT 1409
QY 546 LeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArg 565
Db 1410 GTTGTGTCGGTGTGATGCACCAATCTACTTTGCTAACATAAGTTACATAAAGGACAGG 1469
QY 566 IleThrArgTrpIleAspGluGluGluGluArgIleLysAlaThrGlyGluThrSerLeu 585
Db 1470 TTGCGTGATGATGAGCTCAAACTCCCAATTCAAACCGTGGACCTGATGTTGGAAGGGTG 1529
QY 586 GlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMet 605
Db 1530 TACTTTGTGATCTCGAGATGTCCTCTTTACATACATGACCTCGAGCGCTGTTCAAGCT 1589
QY 606 LeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnPro 625
Db 1590 CTCAAGACCTGCACCAAGAATAACAAGCAGCGGCATCCAGATTGCTATAGCGAATCCT 1649
QY 626 ValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLys 645
Db 1650 AACCGGAGGTGCACCTATTGCTGTCAAGAGCGGGCATCATCGACATGTTGGCGCAGGG 1709
QY 646 TrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSer 665
Db 1710 TGGTGTTCGTCGAGTGCACGACGCGGTGCAAGTATGCTCCAGCATGTGCGGAGTTCA 1769
QY 666 LysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db 1770 TCGTCGAATGCCAATTAAGTTATCCCCACACAGCGGCTGTGGGAACCTG 1814
RESULT 7
US-09-720-317A-7
; Sequence 7, Application US/09720317A
; Patent No. 6896292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Helianthus tuberosus
US-09-720-317A-7
Alignment Scores:
Pred. No.: 9,44e-81 Length: 780
Score: 731.50 Matches: 145
Percent Similarity: 79.64% Conservative: 31
Best Local Similarity: 65.61% Mismatches: 38
Query Match: 21.05% Indels: 7
DB: 4 Gaps: 1
US-10-762-049-18 (1-680) x US-09-720-317A-7 (1-780)
QY 467 AlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyr 486
Db 17 GCGCGCAATCACCCTCTGGACACTAGACAAATTCGACTTTGTGTATGCGATGATGCGATAC 76
QY 487 IleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerVal 506
Db 77 TTTGGTGTGTCTTGGAGAGTGTGAATTTGGATTAGTTATCGCGTCCGATGTCGTG 136
QY 507 LeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsn 526
```

```
Db 137 CTTAGGTACTCTTATTTGCTCGAGGCCAAGAACATCGACGCTAGGTCTCATACCCGAT 196
QY 527 SerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeu 546
Db 197 TCCATATTATAGAAATATGATCAATACCAAAATGCGAAAGCGTTCCAGAAATCTTG 256
QY 547 IleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIle 566
Db 257 ATACTTCAATCGAAGCACCTATTACTTTGCTAACTCTAGCTACTTTGAGGGAAGGATT 316
QY 567 ThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGln 586
Db 317 GTGAGATGGTGTGATGAGGAGGAGATAGTGTGAAGTCTTTAAAGGAGAAATGACTTGCAA 376
QY 587 TyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeu 606
Db 377 TAGTCATCTTTCATTTAGTGTCTGTGGAATATTTATACAAAGTGGGATACAAATGCTT 436
QY 607 GluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProVal 626
Db 437 GGAGAAGTTAAAGAGTTATGGAAGAGAGGGCTAAAGTTGTTTAGCGAATCCGGGC 496
QY 627 SerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrp 646
Db 497 GGAGAGGTAAATAAAGAGATGAACAAAGCGAAGTTTCATAGAGGTGATCGGGCAAGAAATGG 556
QY 647 IleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerLys 666
Db 557 ATATATCTAACAGTGGGAGACGGTTGAGCGTGCACCTTTATGCTTCATACTTACAAAG 616
QY 667 ThrAsnProLysLys-----AspGluThrGluGlyTrpAsnAsn 679
Db 617 AACGCGGAAACCCCACTTCTGATCAGATCAGGAAAGAGAGTCCGAAACGACATATAT 676
QY 680 Val 680
Db 677 GTC 679
RESULT 8
US-09-252-991A-7241
; Sequence 7241, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7241
; LENGTH: 1791
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7241
Alignment Scores:
Pred. No.: 1.27e-74 Length: 1791
Score: 687.50 Matches: 175
Percent Similarity: 49.68% Conservative: 118
Best Local Similarity: 29.68% Mismatches: 246
Query Match: 19.78% Indels: 51
DB: 4 Gaps: 14
US-10-762-049-18 (1-680) x US-09-252-991A-7241 (1-1791)
QY 91 ProIlePheGluTrpAlaPro-----LysTyrThrPheGlnPheLeu 104
Db 91 LysTyrThrPheGlnPheLeu 104
```

Db	73	CCACTCGCCCGCTGGGTACCCGGAGCTCGATAGCTGCTCTCACTACCGCGCGCGCTGGTTC	132
Qy	105	LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer	124
Db	133	CGTCCGGATGTCACAGGCGGGCTGTCGTAGCGCAATCCAGATCCCAACCGCCATCGCT	192
Qy	125	TyrAlaIysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIleProPro	144
Db	193	TACGCCCAGATCGCGGCTTCCCGCCAGGTAGCGCTCTATGGTGTGCATCTCCTCCGATG	252
Qy	145	LeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGly	164
Db	253	CTGATCTACGCCCTTCATCGGCAGCTCGCGCAGTTGATGTAGGCCCGCAGCGCGCG	309
Qy	165	SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspPro-----AsnGluAspPro	182
Db	310	-----ACCGCGCGGATGGTCGCGCGCCATCAACCCCGCTGGCGGGCGGATCCG	360
Qy	183	LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla	202
Db	361	CAGCGCTGTGTGGACCTGCTCATGATCGTCGCGATCATGTGTCGGCTGTCTTCATCGC	420
Qy	203	LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGly	222
Db	421	GCCGGCTTGGCCCGCGCGGTTTCATCGCCAGCTTCTCTCGCACCGATCTCGTGGCG	480
Qy	223	PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuIysSerIleLeuGlyLeu	242
Db	481	TACCTCAACGGCATCGGCTGCTGTCGGGCACTGGGCACAGCTGTTGCGGTAC	540
Qy	243	GluHisPheThrHisGly-----AlaAspIleIleSerValMetArgSerValPheThr	260
Db	541	GAGCGCGCACACAGCGCTTCGTCGCGGATCTCTCGCGCTGCTGGAGAACCTGCTG	597
Qy	261	GlnThrHisGluTyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeu	280
Db	598	-----CACATCCATTCGCCACGCTGATCTTCGGCAGCCTCTCGCTTGTCTGATGA	648
Qy	281	LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTyrValSerAlaMetAla	300
Db	649	GTGCTGTG-----CCGCGCGCGCTTC-----CGCAGTTGCCCGGG	684
Qy	301	ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis	320
Db	685	CGCTCTCGCGCGTCTCTCGCCAGCTCGCGCGCGCTGCTGGCGCTGGATCGGTAC	744
Qy	321	GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeu-----	337
Db	745	GGCGTGAACTGCTGTGCGAGGTACCGCGCGGCTG-----CCGCAACTGAGCTGGCCG	798
Qy	338	---ThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleVal	356
Db	799	CAGACAGCTGGAGGAACCTGCTGGCGCAGCC-----ACCGGATCAGC	852
Qy	357	ValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLys	376
Db	853	GTG-----GTCAGCTTCTGCAGCGCATGTCTACCGCGCGCAGCTTCGCGCCCGTCAC	906
Qy	377	AsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGly	396
Db	907	GGCTACAGCATCAACCCCAACACAGAAATCGTCGCCCTCGGCGCTGGCCACATCGCGCC	966
Qy	397	SerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsn	416
Db	967	GGGCTCTCCAGGGCTTCGCCCATCGCGCGCGCGCACTCAGCACCGCGGTGAACGACATG	1026
Qy	417	AlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeu	436
Db	1027	GTGCGCGCAGACCCCAACTGGTCGGGTGTCTCGCGCTGGTGTATCGCGCCACCTG	1086
Qy	437	LeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleVal	456
Db	1087	TTGCTGTGAACAGCCCTGGGTGCGGATCCGATGCGCGCTGCGGTGCGGTCTGCTGTG	1146

Qy	457	SerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPhelysValAspLys	476
Db	1147	CTGGCCGGCTGGGGCTGTGATCGAGCTGCAAGCGCTGAAGGGCTTCTTGGAAAGCTCAGCCGC	1206
Qy	477	PheAspPheValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIle	496
Db	1207	TTGAGTTTCAGCTCTGCTGCCCTGCTGACACCGTGGCGTGTCTACGCTGCGGTGTGCTCCG	1266
Qy	497	GlyLeuValIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgPro	516
Db	1267	GGAATCTTCGTCGCGCTGACATCGCGTACTGCGCTGTCTTACTACACTATCGCCCG	1326
Qy	517	ArgThrPheValLeuGly-----AsnIleProAsnSerValIleTyrArgAsnVal	533
Db	1327	AGCGACCGCTGCTCGGCTCGATCGAGCGCATCGAGCGCCAGGTC-----GAGCTG	1377
Qy	534	GluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaPro	553
Db	1378	GCCAAGTACCGGAGGCCACCACCTCTTCGCGGCTGGTGATCTACCGCTTCGACGCGCCC	1437
Qy	554	IleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGluGlu	573
Db	1438	CTGCTGTCTTCAACGCCGAGTACTTCAAGCAGCGGTGCTTCCCGCTGGTGGAC-----	1491
Qy	574	GluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSer	593
Db	1492	-----GGCAGCGAGCGCGCGAATGCGGTGCTGTCTCAACGCGGAA	1530
Qy	594	AlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluValLysIleThr	613
Db	1531	GCCATGACCACTCGACATCAGCGGCTGGCCACCTTCACAGGTGCAACAGATCCTC	1590
Qy	614	GluArgArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLysLysLeu	633
Db	1591	AAGGCCAGCGCGTGCACCTGTGCTGGCGCGGTGCACCGGCAGACGCTGGACCTGCTG	1650
Qy	634	AsnLysSerLysPheGlnAsnHisLeuGlyLysTrpIleTyrLeuThrValGluGlu	653
Db	1651	CACGCTCGAGATGCTGGCGGAGATCAAGCCCGCTGTGTTCAGTTCGGTGGCTCC	1710
Qy	654	AlaValGlyAlaCysAsnPheAsnLeuArg	663
Db	1711	GGGTCAGTGCCTATCGCTACTGGCTGGCG	1740
RESULT 9			
US-09-252-991A-7184			
; Sequence 7184, Application US/09252991A			
; Patent No. 6551795			
; GENERAL INFORMATION:			
; APPLICANT: Marc J. Rubenfield et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDO			
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.136			
; CURRENT APPLICATION NUMBER: US/09/252,991A			
; PRIOR FILING DATE: 1999-02-18			
; PRIOR APPLICATION NUMBER: US 60/074,788			
; PRIOR FILING DATE: 1998-02-18			
; PRIOR APPLICATION NUMBER: US 60/094,190			
; PRIOR FILING DATE: 1998-07-27			
; NUMBER OF SEQ ID NOS: 33142			
; SEQ ID NO 7184			
; LENGTH: 1767			
; TYPE: DNA			
; ORGANISM: Pseudomonas aeruginosa			
US-09-252-991A-7184			
Alignment Scores:			
Pred. No.:		2,91e-73	Length: 1767
Score:		676.50	Matches: 172
Percent Similarity:		50.17%	Conservative: 117
Best Local Similarity:		29.86%	Mismatches: 242
Query Match:		19.47%	Indels: 45

## RESULTS

```

US-09-252-991A-7184
; Sequence 7184. Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO AC
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTIC
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7184
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: pseudomonas aeruginosa
US-09-252-991A-7184

```

Alignment Scores:	
Pred. No.:	2.91e-73
Score:	676.50
Length:	1767
Matches:	172
Percent Similarity:	50.17%
Conservative:	117
Best Local Similarity:	29.96%
Mismatches:	242
Query Match:	19.47%
Indels:	45

```
DB: 4 Gaps: 13
US-10-762-049-18 (1-680) x US-09-252-991A-7184 (1-1767)
Qy 99 TyrThrPheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAla 118
Db 11 TACCGCGCGCCCTGGTTCGCGATGTCACGCGCGGCTGTCGTAGCGCAATCCAG 70
Qy 119 IleProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyr 138
Db 71 ATCCCAACCGCCATCGTTACGCCAGATCGCGGCTTCCGCGCCAGGTAGGCTCTAT 130
Qy 139 SerSerPheIleProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaVal 158
Db 131 CGGTGCATCTCGCGATGCTGATCTACGCCCTGATCGCGAGCTCGCGGAGTTGATGTA 190
Qy 159 GlyThrValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspPro 178
Db 191 GGCCCCGACGCGCGC-----ACGCGCGCATGTGTCGCCGCCCATCACCCCG 238
Qy 179 -----AsnGluAspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAla 196
Db 239 CTGCGCGCGCGCATCGACGCCCTGGTGACCTGTGATGATCGTCGCGATCATGGTC 298
Qy 197 GlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSer 216
Db 299 GGCGCTTCTCCATCGTCGCGCGCTGGCGCGCGGTTTCATCGCCAGCTTCTCTCG 358
Qy 217 HisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeu 236
Db 359 CGACCGATCTGTGTCGGCTACCTCAACGCGCATCGGCTGAGCTGTGTCGGGCAACTG 418
Qy 237 LysSerIleLeuGlyLeuGlnHisPheThrHisGly-----AlaAspIleIleSerVal 254
Db 419 GGCAAGCTGTTGGCTACGAGCGCGCGCACGAGCGGCTTGTGCGCGCATCTCGCGCTG 478
Qy 255 MetArgSerValPheThrGlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCys 274
Db 479 CTGGAAACCTGTGT-----CACATCCATTGGCGACGCTGATCTCGGCAGC 526
Qy 275 ValPheIlePhePheLeuLeuSerThrArgTyrPheSerIlyLysArgProArgPhePhe 294
Db 527 CTCCTCCTCTGTGTGTGTGTGTGT-----CCGCGCGCGCTTC 565
Qy 295 TrpValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuValTyrPhe 314
Db 566 ----CCGCAGTGGCCGGCGGCTCTCGGGGCTTCTGCGCAGCTCTCGCGGGCGCTG 622
Qy 315 ThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnPro 334
Db 623 CTGGGCTTGGATCGCTACGCGTGGAACTGCTTGGCGAGGTACCGCGCGGCGCTG----- 676
Qy 335 ProSerLeu-----ThrAsnLeuValPheValSerProTyrMetThrThrAla 350
Db 677 CGCAACTGAGTGGCGCAGACACCGCTGGAGAACTGAAAAAGCCCTGTGCGCGACGCG 736
Qy 351 ValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArg 370
Db 737 -----ACCGCATACGGTG-----GTCAGTTCTGCAGCGCCATGCTACCGCGCGC 784
Qy 371 SerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGly 390
Db 785 AGCTTCGCGCGCGTCACGCTACAGCATCAACCCCAACACAGAAATTCGTGCGCTCGGC 844
Qy 391 ThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArg 410
Db 845 CTGGCCAAATCATGCGCGCGGGTCTCCAGGGCTTCGCCATCAGCGCGCGCCACTCACGC 904
Qy 411 SerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeu 430
Db 905 ACCGCGGTGACACATGTCGGCGCGCAGACCCAACTGTCGGCGGTGTCGCGCGCTG 964
Qy 431 AlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValVal 450
```

```
Db 965 GTGATCGCGCGCACCTGTTGCTGTGAACAAGCCCTGGGCTGGGTCGCGATGCGCGG 1024
Qy 451 LeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaIleHis 470
Db 1025 CTCGTGCGGCTCTGTTGCTGCGCGGCTGGGCGCTGATCGACGTGCGCGCTGAAGGC 1084
Qy 471 LeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValVal 490
Db 1085 TTCTGGAAGCTCAGCGGCTTCGAGTTACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTC 1144
Qy 491 PheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeu 510
Db 1145 AGCGTCGCGCTGCTGCGGGAATCTTCGTCGCGCTGAGCATCGCGTACTGCGCTGCTC 1204
Qy 511 LeuPheIleAlaArgProArgThrPheValLeuGly-----AsnIleProAsnSer 527
Db 1205 TACTACACTATCGCGCCGAGCGACGCGGTGCTGCGCTGCGATCGCGCATCGCGCCAG 1264
Qy 528 ValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIle 547
Db 1265 GTC-----GAGCTGGCCAAGTACCGCAGGCGCACCTCTTCCCGGCTGTGTATC 1315
Qy 548 LeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThr 567
Db 1316 TACCGCTTCGACCGCGCTGCTGTTCTTCAACGCGGAGTACTTCAAGCAGCGGTGCTT 1375
Qy 568 ArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyr 587
Db 1376 GCCGTGGTGGAC-----GGCAGCGAGCGCGCGAATGCG 1408
Qy 588 ValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 607
Db 1409 GTGCTGCTCAACGCGCGAGCCATGACCACTCGACATCAGCGGCTGGCCACCTCCAC 1468
Qy 608 GluValLysLysIleThrGluArgGluLeuGlnLeuValLeuValAsnProValSer 627
Db 1469 GAGGTGCAACAGATCTCAAGGCCCGCGGCTGACCTGTCTGCTGCGCGGTGACCGGG 1528
Qy 628 GluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIle 647
Db 1529 CAGACGCTGGACCTGCTGCAAGCTCGAGTATGCTCGCGAGATCAAGCGCGCTGGTG 1588
Qy 648 TyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArg 663
Db 1589 TTCAGCTCGGTGCGTCCGGGTGAGTGCTATCGTACTGCTGCTGCTGCTGCTGCTG 1636

RESULT 10
US-09-720-317A-5
; Sequence 5, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Zea mays
US-09-720-317A-5

Alignment Scores: 2,05e-73 Length: 1240
Pred. No.: 675.50 Matches: 132
Score: 67.22% Conservative: 69
```

```
Best Local Similarity: 44.15% Mismatches: 97
Query Match: 19.44% Indels: 1
DB: 4 Gaps: 1

US-10-762-049-18 (1-680) x US-09-720-317A-5 (1-1240)

QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
DB 6 AGCTTCACGGAGGATCGCGTGGCGGATCTTCGCCCTCCGTAAGAGGATCAGACTC 65
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValAlaGlySerPheThrSer 400
DB 66 GACGCAACAGGAGATGCTGCCCATGGGTTCTCCCAACGTTGCTGGTCTCTGTCTCTCG 125
QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
DB 126 TGCTATGTGCAACAGGTTCTGCTCCGCAACGCGAGTGAACCTTCAGCGGGGGCCAGG 185
QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuPheLeuThr 440
DB 186 TCGACGTTTCAACATCGTATGTCATACCGTGTTCGTCACCTGGAGCTGTTTCATG 245
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
DB 246 AAGCTCTCTACTACAGCCCATCGCGTGTCTGCCCTCCATCATCTCGCTCTTCGCG 305
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
DB 306 GGACTGATCGACATCAAGAGGCGTGCAGCATATGGAAGATCGAAGATGGATTTCCCTC 365
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
DB 366 ACCTGCTCGGTGTGTGGTGTCTGTTGGTGGTGGAGATGGGCTTTCAGATT 425
QY 501 AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal 520
DB 426 GCACTTGCACTTCCTTCGCAAGATCATCATACAGTCGCTTCGGCTCAGGTGGAGATC 485
QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
DB 486 CTTGGCAGGCTTACAAGGAGCAGATATCTTCTGCAGCGTCAGGAGTACCCTGTAGCTGC 545
QY 541 HisValProGlyMetLeuIleLeuGluIleAspAlaPro---IleTyrPheAlaAsnAla 559
DB 546 CTAACCTCGACTGCTACTGCTATACGCGTCGACACATCTTCCTCTCTCATCAACGCC 605
QY 560 SerTyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAla 579
DB 606 ACTTCCTCAAGAAAGGATCAGAGTGGTGGGAGGAGTGGAGACTCAATGGA 665
QY 580 ThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAsp 599
DB 666 AAGCGGAGGAGAGATACAAAGCAGTGTCTCTGATATGCAAGTGTGGTAAACATCGAC 725
QY 600 ThrSerGlyIleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGln 619
DB 726 ACTTCAGACTCACTGCACTGGAAGAATAACAAAGAGTGGTGTCTCTTGGCTTACAG 785
QY 620 LeuValLeuValAsnProValSerGluValMetLysLysValLeuAsnLysSerLysPheGln 639
DB 786 ATGCTATAGCCATCGCGGTGAAGGAGGATTCAGAAAGATGAAGTGTACAGGTGGTG 845
QY 640 AsnHisLeuGlyLysIleTyrIleThrValGluGluAlaValAlaGlyAlaCys 658
DB 846 GACAGGGTAGCAGGAGTGGATCTTCATGACAGTAGGTGAAGCGGTGGAGGCGCTGT 902

RESULT 11
US-09-785-381-4
; Sequence 4, Application US/09785381
; Patent No. 6602992
; GENERAL INFORMATION:
; APPLICANT: DALLOS, Peter
; APPLICANT: ZHENG, Jing
```

```
; APPLICANT: MADISON, Laird
; TITLE OF INVENTION: A MAMMALIAN PRESTIN
; FILE REFERENCE: 0290-3701
; CURRENT APPLICATION NUMBER: US/09/785,381
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,461
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2441
; TYPE: DNA
; ORGANISM: Mus sp.
; US-09-785-381-4

Alignment Scores:
Pred. No.: 7,8e-70 Length: 2441
Score: 651.00 Matches: 191
Percent Similarity: 45.95% Conservative: 144
Best Local Similarity: 26.20% Mismatches: 270
Query Match: 18.73% Indels: 124
DB: 21 Gaps: 21

US-10-762-049-18 (1-680) x US-09-785-381-4 (1-2441)

QY 35 MetAsnAsnPheGluArgValHisGlnValGluValProPro----- 48
DB 198 CTCGACAGCATGATCATGCTGAAGAAATGAATAATCCCTGCAGACCCAGAGGTACTAC 257
QY 49 ---ProGlnProPhePhe-----LysSerLeuLysTyrSerLeuLys 61
DB 258 GTGGAAGGCCCATCTTTCAGTCATCTCTGCTCCAGAGAGGCTGCACGTCAAGGACAAA 317
QY 62 GluThrPhePheProAspAspProLeuArgGln---PheLysAsnLysProAlaSerLys 80
DB 318 GTCACAGAGTCCATTGGAGATAAGCTGAAGCAGCATTCACGTGTACTCCT-----AAA 371
QY 81 LysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThr 100
DB 372 AAAATAAGAAACATCATTTACATGTTCTCTGCTTATCATTAAGTGGCTGCCAGCATATAA 431
QY 101 Phe---GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIle 119
DB 432 TTCAGAGGATGTGTGTAGTGACTTGTCTCGGCAATAGCACTGGGGTACTCCACTT 491
QY 120 ProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSer 139
DB 492 CCCCAAGGCTTAGCCTTCGCCATGCTGGCAGCGCTCCCTCCGGTGTGGCTGTACTCA 551
QY 140 SerPheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGly 159
DB 552 TCGTTTTTACCCCGTTATCATGTACTGTCTTTTGGAAACCTCAAGACACATATCTATAGT 611
QY 160 ThrValAlaValGlySerLeuLeuMetGlySerMetLeu----- 172
DB 612 CCTTTTGTGTTATTAGTTGATGATGGAGGTGGCCCTCGCTGTAGTACCAGATGAT 671
QY 173 -----SerAsnAlaValAspProAsnGlu---AspProLysLeuTyr 185
DB 672 ATTGTATCCAGGAGGAGTAAATGCAACCAACGGGACAGAGCCAGAGATGCCTAAGA 731
QY 186 LeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeu 205
DB 732 GTGAAAGTCGCCATGTCTGTACTTCTTACCTTTTACGGAATCATCAGTTTTCCTAGGTGC 791
QY 206 PheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGly 225
DB 792 TGATGTTGGATTGTGGCCATATACCTCAGGAGCCATTGGTGGAGGCTTTTACCCT 851
QY 226 GlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPhe 245
DB 852 GCGGCTGCTGCCACGTGTTTACGTCCATGTTTAAATACCTGTTTGGGGTCAAAACAAG 911
```









[illegible]

```

/ AFFILIATION: Wiegand, Roger C.
/ TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
/ FILE REFERENCE: 38-10(15849)B
/ CURRENT APPLICATION NUMBER: US/09/902,540
/ CURRENT FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: 60/217,883

```

```

; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 768
; LENGTH: 4854
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902,540-768

```

**Alignment Scores:**

Pred. No.:	2,318-68	Length:	4854
Score:	643.50	Matches:	162
Percent Similarity:	50.44%	Conservative:	124
Best Local Similarity:	28.57%	Mismatches:	256
Query Match:	18.53%	Indels:	25
DB:	4	Gaps:	7

US-10-762-049-18 (1-680) x US-09-902-540-768 (1-4854)

Qy	91	ProilePheGluTrpAlaProLysTyThrPheGlnPheLeuLysAlaAapLeuIleAla	111
Db	68	CGGGGCTCGGGCAGCAGCGCGCTATCAGCGCCGCTGGTTCCGGGGGAGACCTGCTCTCG	127
Qy	111	GlyIleThrIleAlaSerLeuAlaIleProGlnGlnIleSerTyAlaLysLeuAlaAaSn	130
Db	128	CGCTCACCATTGGTGGATGCTATCCCCAGGGCTGGCTACGGCGAGCTGTGGGC	187
Qy	131	LeuProProIleLeuGlyLeuTySerSerPheIleProProLeuIleTyAlaMetMet	150
Db	188	GTGCGTCCAGCAGCGGGCCTCTACGCGGGCGTGTGGGAATGCTGGCTATGGCGTCTC	247
Qy	151	GlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMetGlySer	170
Db	248	GGGCGCTCGGGCACCCTCATCTCGAGCCCGAAGAGGGGTGCGGCATCTTCACGCGCGCG	307
Qy	171	MetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyIleLeuAlaPhe	190
Db	308	CGCTCGCG---CGGTGGCGGGGGGGCGCCCGCGCGCTACGCTGGCGGCA	364
Qy	191	ThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuLeuGlyIleuPheArgLeuGlyLeu	210
Db	365	CTGCTGGCGCTGTGTGGCGTGTGAGCCCTGTCTCGAGGGGTGCTCAAGGTGGGCGCG	424
Qy	211	IleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrVal	230
Db	425	CTCGGAGATTCTGTCTCAAGCCCATCTCATCGCTACATCAATGGCGGGCGCTCATC	484
Qy	231	ValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis-----PheThrHisGly	248
Db	485	ATCATCGGCAGCAGCTCGCGGCTCTTCGGCGAGGAGCGCCAGTCCGACACGTTCTCG	544
Qy	249	AlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrpGlu	268
Db	545	GGTCAAGTGTTCGAGGTGGCCACCCACCTC---GAACGGACCCACGTC-----CCG	592
Qy	269	SerAlaValLeuGlyCysValPheIlePheLeuLeuSerThrArgTyPheSerLys	288
Db	593	ACGCTCTGCTGGGCGTGGGCGTCATACCGCGCTGTGTGTGTCGGCGCAACTCTGCCCC	652
Qy	289	LysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGly	308
Db	653	NAGGTGCCC-----GGTCCGCTCATCTCGTGGTGGTCTCACC	688
Qy	309	SerLeuLeuValTyPheThrHisAlaGluLysHisGlyValGluValIleGlyGluLeu	328
Db	689	ACGGTGGCGGGAGGGCTGTTCCAGTTGTGAGCATGGGGGCATCAAGGTGTGGGCGCCCTC	748
Qy	329	LysLysGlyLeuAanProSerLeuThrAanLeuValPheValSerProTyfMetThr	348
Db	749	GCGCCGAGCCCCCGCTCCAGGCTCGCGTCCCTCGCTTCGAGGAC-----GTGCGG	802
Qy	349	ThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaVal	368
Db	803	TCGCTGCTCCCGCGCGCTTACGCTGGCGCTGTCTCACTACGCGCAGCTCGTGTGACG	862

RESULT 15

```

US-09-949-016-4516
; Sequence 4516, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN
; TITLE OF INVENTION: WITH HUMAN DISEASE
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

```

[illegible]

```
QY 569 -----TrpIleAspGluGluGluGluArgIleLysAlaThrGlyGluThr--- 583
Db 1852 ATAAAGTGGCTTGGAGAGGAGCAAGAGAGAGATCAAGAAAAAGTAGTCACTCTT 1911
QY 584 -----SerLeuGlnTyr 587
Db 1912 GGTGGATCCAGGATGAAATGTGAGTCAACTTTCCCATGATCCCTTGGAGCTGCATACT 1971
QY 588 ValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 607
Db 1972 ATAGTGATTGACTGCAATTCATTTTATAGATACAGCAGGGATCCACACACTGAAA 2031
QY 608 GluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProValSer 627
Db 2032 GAAAGTTGCGCAGAGATTATGAAGCCATTGGAATCCAGGTTCTGCTGGCTCAGTGCATCCC 2091
QY 628 GluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIle 647
Db 2092 ACTGTGAGGGATTCCCTAACCAAGGAGAGATATTGCAAAAAGGAAGAAACCTTCTC 2151
QY 648 TyrLeuThrValGluGluAlaVal 655
Db 2152 TTCTATAGTGTGTATGAAGCGATG 2175
```

Search completed: August 31, 2005, 16:06:59  
Job time : 386 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 31, 2005, 11:53:46 ; Search time 1127 Seconds  
(without alignments)  
3947.916 Million cell updates/sec

Title: US-10-762-049-18  
Perfect score: 3475  
Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGNNV 680

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Xgapop 10.0 , Xgapext 0.5

Xgapop 6.0 , Xgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10762049/runat\_3082005\_121246\_6030/app\_query.fasta\_1.839  
-DB=Published Applications NA -QFMT=fastp -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCUT=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10762049 @CGN 1 1 480 @runat\_3082005\_121246\_6030  
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:\*\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
23: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
27: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2595.5	74.7	1977	9	US-09-938-842A-1502
2	2595.5	74.7	1977	11	US-09-938-842A-1502
3	2330.5	67.1	2372	18	US-10-437-963-38184
4	2036	58.6	2008	18	US-10-424-599-131116
5	1856	53.4	2727	20	US-10-425-115-50601
6	1845.5	53.1	2651	20	US-10-425-115-182207
7	1844.5	53.1	2651	20	US-10-425-115-182207
8	1773.5	51.0	2780	19	US-10-437-963-77005
9	1773.5	51.0	2780	19	US-10-437-963-77005
10	1762	50.7	2656	18	US-10-425-114-15719
11	1762	50.7	2656	18	US-10-425-114-15719
12	1734.5	49.9	2429	19	US-10-437-963-73410
13	1698	48.9	4390	19	US-10-437-963-77012
14	1676.5	48.2	1518	18	US-10-424-599-131117
15	1563.5	45.0	2251	19	US-10-437-963-72062
16	1513.5	43.6	1980	17	US-10-260-238-1066
17	1500	43.2	2336	18	US-10-424-599-52175
18	1460	42.0	1992	18	US-10-425-114-13333
19	1446.5	41.6	3252	19	US-10-437-963-78357
20	1436	41.3	3201	20	US-10-425-115-44188
21	1406	40.5	2242	20	US-10-425-115-40196
22	1379.5	39.7	2506	18	US-10-424-599-74393
23	1218	35.1	2156	18	US-10-424-599-60281
24	1215	35.0	2156	20	US-10-739-930-3837
25	1204	34.6	1809	18	US-10-424-599-83137
26	1185.5	34.1	1632	18	US-10-424-599-74389
27	1175.5	33.8	2913	19	US-10-437-963-76453
28	1134.5	32.6	2346	20	US-10-425-115-22539
29	1016	29.2	1303	18	US-10-424-599-139734
30	1010.5	29.1	1551	18	US-10-424-599-54350
31	1009	29.0	1247	18	US-10-424-599-130951
32	1002	28.8	2058	9	US-09-938-842A-717
33	1002	28.8	2058	11	US-09-938-842A-717
34	955.5	27.5	2988	18	US-10-437-963-92700
35	934	26.9	680	18	US-10-424-599-32307
36	923	26.6	1019	19	US-10-767-701-11661
37	901	25.9	1174	20	US-10-425-115-75765
38	900	25.9	536	18	US-10-424-599-32308
39	898	25.8	674	19	US-10-767-701-701
40	892.5	25.7	1115	19	US-10-767-701-9435
41	887	25.5	1069	20	US-10-425-115-3619
42	814	23.4	1802	19	US-10-437-963-7407
43	737.5	21.2	1018	18	US-10-424-599-18417
44	736.5	21.2	841	18	US-10-424-599-135930
45	715	20.6	788	20	US-10-425-115-75767

# ALIGNMENTS

## RESULT 1

US-09-938-842A-1502  
; Sequence 1502, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938, 842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24

```
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1502
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1502

Alignment Scores:
Pred. No.:      5,91e-263      Length:      1977
Score:          2595.50      Matches:      498
Percent Similarity: 87.73%      Conservative: 81
Best Local Similarity: 75.45%      Mismatches:  77
Query Match:      74.69%      Indels:       5
DB:               9          Gaps:         3

US-10-762-049-18 (1-680) x US-09-938-842A-1502 (1-1977)

QY 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1 ATGGGACGGAGGACTACATTCCTCAAGGACGGAGGAGCTTCACGCCGCCACCCAC 60

QY 44 ---ValGluValProProGlnProPhePheLeuSerLeuLysTyrSerLeuLysGlu 62
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 61 ACGGTGGAGGCTCCACAACCTCAACGGTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120

QY 63 ThrPhePheProAspAppProLeuArgGlnPheLeuAsnLysProAlaSerLysLysPhe 82
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 121 ACTCTGTCTTCACAGCAGCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT 180

QY 83 MetLeuGlyLeuGlnPhePheProIlePheGluTyrAlaProLysTyrThrPheGln 102
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 181 GTGTTAGGCTCAAACTCTCCCGATTTCCGATTTTCGATGGGCACCACTACAAATCTCAAG 240

QY 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 241 TTCTTCAATACAGATCTCATCGCGCGGAATCACCATCGTAGTCTCGCCATCCCTCAGGCG 300

QY 123 IleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIle 142
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 301 ATCAGTTACGCCAAACTCTGCTAACTTGCCTCCCAATTTCTTGGCCCTTTATTCGAGTTTGT 360

QY 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 361 CCGCCATTTGGTATACCGGCTGCTAGGAGTTCAAGGACTTAGCGGTGGAAACGGTTGCG 420

QY 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 421 GTTGGCTCTCTGTGACAGTGGGTGCTGAGCAAAAGAGTTGATGCTGAGAAAGATCCT 480

QY 183 LysLeuTyrLeuHisIleuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 481 AAGCTTTACCTTCACCTTGTCTTTCACGCCCACTTTTTCGCCGCGCTTCTCGAAGCCTCT 540

QY 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleG 222
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 541 CTTGGAAATTTTCAAGGTTAGGGTTCAATAGTGACTTTTCTATCGCATGCAACGATAGTAGG 600

QY 222 yPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 601 ATTCATGGGAGGACGACGACGCTGCTGAGTCTGCAACAGCTTAAGGGTATTTTCGGACT 660

QY 242 uGluHisPheThrHisGlyAlaAspIleLeuSerValMetArgSerValPheThrGlnTh 262
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 661 TAAACATTTTCACAGACTCTACCGATGTATCTCTGTCATGCGGTTCGGTTTCTCCCAAC 720

QY 262 rHisGluTyrArgTyrPheGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 721 TCACGAGTGGAGATGGGAAAGTGGCGTTCTTGGATGGTTTTCTATTCTTCTCTC 780
```

```
QY 282 rThrArgTyrPheSerLysLeuArgProArgPhePheThrValSerAlaMetAlaProLe 302
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 781 CACCAGATATTTGAGCATCAGAAACCAAAATTCCTTTGGTGGCGCGATGCTCTCTTT 840

QY 302 uThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVa 322
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 841 GACCTCAGTGATTCCTTGGAAAGTCTCTTGGTTACTTCACTCAGCTGAGAGACATGGTGT 900

QY 322 lGluValIleGlyGluLeuLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 901 TCAAGTGATAGGGACCTGAGAAAGGGTTGAATCCACTCTCGGTTCTGATCATCTT 960

QY 342 eValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleLeuSerLe 362
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 961 TACTTCCCCTTACATGTCACAGCTGTCAAAACTGGCCTCATCACTGGCATCATCTGCTCT 1020

QY 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGl 382
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1021 CGC-GAAGGAGTAGCAGTGGGAGGAGTTTTCGATGTTTCAAGAACTCAACATAGACGG 1079

QY 382 yAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr 402
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1080 GAACAAAGAGATGATAGCGTTTGGATGATGAACATCGTTGGTCTCTTCCATCTTGTTA 1139

QY 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1140 CCTCAACACCGGACCATTTTCAAGTCGGCAGTGAACATACAACGGGGTTGCAAGACCGC 1199

QY 422 alaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLe 442
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1200 AATGTCCAACATAGTAGTGGGATTCGGGTTATGTTACACTCTCTTCTCACCCGCT 1259

QY 442 uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1260 TTTTCTACACACACACCTCTCTCTGCCATCATCATATTCGGCAATGCTCGGACT 1319

QY 462 uIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCy 482
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1320 CATTCGACTATCAAGCTGCCATCCATCTCTGGAAAGTTGACAAAGTTCGACTTCTCTGCTG 1379

QY 482 sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI 502
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1380 CATGACGCTACGTGGGGTCTATTCGGCAGTGTAGAGATTGGACTCGTCTGACCGGT 1439

QY 502 eValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuGl 522
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1440 GCGGATATCTATAGCGAGGTGTTGCTGTTGTGTGCGAGCCCAAAACTGCGGTGAGGG 1499

QY 522 yAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa 542
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1500 AAACATACCAACACGATGATCTATAGGAACACTGAGCAGTAGTACCCATCATCAAGAACCCT 1559

QY 542 lProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrIle 562
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1560 TCCTGGTATTCATCTTGGAGATTGATGCTCCCATCTACTTGTGTAATGTCAGTTACTT 1619

QY 562 uArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThrGlyGl 582
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1620 GCGTGAAGAATCATAGGTGGATTGATGAAGAGAGAGAGAGATTAAACAATCAGGAGA 1679

QY 582 uThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGl 602
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1680 GAGCAGCTTACAATATATTATCTCGATATGTCAGCTGTTGGTAAATATATCGACACAACCG 1739

QY 602 yIleSerMetLeuGluValLysIleThrGluArgArgGluLeuGlnLeuValIle 622
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1740 TATTAGCATGATGGTGGAAATTAAGAAAGTCTTACACAGGAGAGCGTTAAAGTTGGTATT 1799

QY 622 uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1800 GTCAATCCAAAGAGAGAGGTGCTGAGAAATTAACACAGATCCAAATTCATCGGTGATCA 1859

QY 641 sLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAs 661
```



```
Db 1860 TTTGGGCAAGAGTGGATGTTCTTAACGGTAGGAGAAGCAGTGGAGGCTTGTAGCTACAT 1919
Qy 661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db 1920 GCTTCACACGCTTTAAACACGACCG---GCCTCAAAAACGAGCGTTTGGAACAACGTA 1974

RESULT 2
US-09-938-842A-1502
; Sequence 1502, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1502
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-1502

Alignment Scores:
Pred. No.: 5,91e-263 Length: 1977
Score: 2595.50 Matches: 498
Percent Similarity: 87.73% Conservative: 81
Best Local Similarity: 75.45% Mismatches: 77
Query Match: 74.69% Indels: 5
DB: 11 Gaps: 3

US-10-762-049-18 (1-680) x US-09-938-842A-1502 (1-1977)
Qy 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
Db 1 ATGGGCACGGAGGACTACACATTCCTCAAGGAGCGGAGGAGCTTCACCGCGCCACCAC 60
Qy 44 ---ValGluValProProGlnProPhePheLysSerLeuLysGlu 62
Db 61 ACGGTGGAGGCTCCACAACTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120
Qy 63 ThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db 121 ACTCTGTTTCAGACGACCTTTTAGACAAATTTAAGAACCAAAATGATCAAGAAATTT 180
Qy 83 MetLeuGlyLeuGlnPhePheProPheGluTrpAlaProLysTyrThrPheGln 102
Db 181 GTGTAGGCTCAAAATCTCTCCCGATTTTCGAATGGCGCACCGCTACAATCTCAAG 240
Qy 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 241 TTCTCAAAATCAGATCTCATCGCGGAATCACCATCGCTAGTCTCGCCATCCCTCAGGGC 300
Qy 123 IleSerTyrAlaLysLeuAlaAsnLeuProPheLeuGlyLeuTyrSerSerPheIle 142
Db 301 ATCAGTTACGCCCAACTTGCTTAATTCGCCCAATTTCTGGCCCTTTATTCAGTTTGTGA 360
Qy 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 361 CCGCCATGTGTATACGGGTGCTAGGGAGTTCGAAGGACTTAGCGGTGGGAACGGTTGCG 420
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
```

```
Db 421 GTTGGCTCTCTGTTGACAGTGCAGTGCAGCAAGAAAGATTGATGCTGAGAAAGATCT 480
Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAa 202
Db 481 AAGCTTTACTCTTACCTTCCACCGCCACCTTTTTCGCCGCGGTTCCTCGAAGCTCT 540
Qy 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleLeuG 222
Db 541 CTTGGAAATTTTCAGGGTTAGGGTTTCATAGTGGACTTTTCTATCGCATGCAACGATAGG 600
Qy 222 yPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242
Db 601 ATTATGGGAGGAGCAGCGGTGTGAGTCTGCAACAGCTTAAGGTATTTTCGGACT 660
Qy 242 uGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnTh 262
Db 661 TAAACATTTTCACAGACTCTACCGATGTATCTGTGTCATGCGTTCCGTTTCTCCCAAC 720
Qy 262 rHisGluTrpArgTTPGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282
Db 721 TCACGAGTGGAGATGGGAAAGTGGCGTCTCTGGATGTGTTTCTATTTCTTCTCTC 780
Qy 282 rThrArgTyrPheSerLysLysArgProArgPhePheTTPValSerAlaMetAlaProLe 302
Db 781 CACGAGATTTTGAGCATCAAGAAACCAAAATTTCTTTGGTGGCGCGGATGCTCTCTT 840
Qy 302 uThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGlyLysHisGlyVa 322
Db 841 GACCTCAGTGATTTCTTGGAAAGTCTCTTGGTTACTTCACTCACGCTGAGAGACATG 900
Qy 322 lGluValIleGlyGluLeuLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342
Db 901 TCAAGTGTATAGGGGACCTGAAGAAAGGTTGAATCCACTCTCCGGTTCTGATCTCATCT 960
Qy 342 eValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLe 362
Db 961 TACTTCCCTTACATGTCCACAGCTGTCAAACTGGCCCTCATCTGCGCATCATTTGCTCT 1020
Qy 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspG 382
Db 1021 CGC-GAAGGAGTAGCAGTGGGAGGAGTTTGGCATGTTTCAAGAACTACAACATAGACGG 1079
Qy 382 yAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTy 402
Db 1080 GAACAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCCTTCACATCTTGT 1139
Qy 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422
Db 1140 CCTCAACACCGGACCAATTTTCAAGGTGGGAGTGAACACTACAACGGGTTTGAAGACCG 1199
Qy 422 aAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuPheLeuThrProLe 442
Db 1200 AATGTCCAACATAGTATGCGGATGCGGTATGTTTACACTCTCTTCTTCACACCGCT 1259
Qy 442 uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462
Db 1260 TTTTCACTACACACCACTCGTCTCTCTCTGTCATCATCATATCCGCAATGCTCGGACT 1319
Qy 462 uIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCy 482
Db 1320 CATTTGACTATCAAGCTGCCATCTCTGGAAGTTGACAAAGTTCTGACTTCTCTCTG 1379
Qy 482 sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI 502
Db 1380 CATGAGCGCTAGTTGGGTCGTATTCGCGAGTGTAGAGATTGGACTCGTCTGAGCGGT 1439
Qy 502 eValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuG 522
Db 1440 GCGCATATCTATAGCGAGGTGTGCTGTTTGTGTCGAGGCCAAAAAATCGCGTGAAGGG 1499
Qy 522 yAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa 542
```

```

Db 1500 AAACATACCAAACAGCATGATCTATAGGAACACTGAGCAGTACCCATCATCAAGAACCGT 1551
Qy 542 lProGlyMetLeulleLeuGluIleAspAlaProIleTyPheAlaIleAsnAlaSerTyLe 562
Db 1560 TCCTGGTATTCTCATCTTGGAGATTGATGCTCCCATCTACTTTGCTTAATGCCAGTTACTT 1619
Qy 562 uArgGluArgIleThrArgTrpIleAspGluGluGluGluArgIleLysAlaThrGlyGl 582
Db 1620 GCGTGAAGAAGATCATAAAGTGGATTGATGAAGAGAGAAAGAGAGAGCTTAAACAATCAGGAGA 1679
Qy 582 uThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGl 602
Db 1680 GAGCAGCTTACCAATATATATTATCTCGATATGTCTGAGCTGTGTGGTAAATATCGACACAACGG 1739
Qy 602 YIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLe 622
Db 1740 TATTAGCATGATGGTGAATATTAAAGAAGTCATTGACAGGAGACGCTTAAAGTTGGTATT 1799
Qy 622 uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe--GlnAsnHi 641
Db 1800 GTCAATCCAAACGAGAGGTCGTGAAGAAATAAACCAATCCAAATTCATCGGTGATCA 1859
Qy 641 sLeuGlyLysLysTrpIleTyLeuThrValGluGluAlaValGlyAlaCysAsnPheAs 661
Db 1860 TTTGGCAAGAGTGGATGTTCTTAAACGGTAGGAGAAGCAGTCGAGGCTTGTAGTACAT 1919
Qy 661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db 1920 GCTTCACAGCTTAAACCGAACCG--GCCTCCAAAAACAGCGCTTGGAAACAACGTA 1974

RESULT 3
US-10-437-963-38184
; Sequence 38184, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 38184
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_41845C.1
US-10-437-963-38184

Alignment Scores:
Pred. No.: 7,65e-235 Length: 2372
Score: 2330.50 Matches: 441
Percent Similarity: 83.84% Conservative: 83
Best Local Similarity: 70.56% Mismatches: 100
Query Match: 67.06% Indels: 1
DB: 19 Gaps: 1

US-10-762-049-18 (1-680) x US-10-437-963-38184 (1-2372)

Qy 43 GlnValGluValProProGlnProPhePheLysSerLeuLysTySerLeuLysGlu 62
Db 149 AGGTGGCGATGCCGGCGGCAACCGCTTCTGGAGACGCTGGGGGGGAACATGAAGAG 208
Qy 63 ThrPhePheProAspAspProLeuArgGlnPhe--LysAsnLysProAlaSerLysLys 81

```

[illegible]





```
Db 1623 GCATGATGGGAGCAATTTTGGAGTCATATTTTCATCGGTGAGATATGCTTGCATTT 1682
Qy 501 AlaileValileSerValLeuArgValLeuLeuPheilleAlaArgProArgThrPheVal 520
Db 1683 GCGGTTGCAATATCTCTTGTAATAATCTCTCCAGTACACGGCCAGAACAGTTTAA 1742
Qy 521 LeuGlyAsnilleProAsnSerValilleTyrArgAsnValGluHistiTyrGlnAsnAlaLys 540
Db 1743 CTTGGAAACCTCTCCAGCAACACTATATACAGAAATGTGAACAGATCTCTGATGTACC 1802
Qy 541 HisValProGlyMetLeuileLeuGluileAspAlaProileTyrPheAlaAsnAlaSer 560
Db 1803 AAGGTTCCAGGGGTGCTGATGTTAGAGTGAGCTCAGCTATATCTACAAACTCTAAC 1862
Qy 561 TyrLeuArgGluArgileThrArgTTPilleAspGluGluGluArgGluArgGluLeu 580
Db 1863 TATGTTAAAGAGAGAACTCTGAGTGGCTAAGAGATGAGGAGGAGCAACACAGGACCAG 1922
Qy 581 GlyGluThrSerLeuGlnTyrValilleleAspMetSerAlaValGlyAsnilleAspThr 600
Db 1923 AAGTTAAACAAAACCTGAGTTTCTAAATGTTGACCTGTCTCTGTAATTGATATCGACACA 1982
Qy 601 SerGlyIleSerMetLeuGluValLysLysLysLysLysLysLysLysLysLysLys 620
Db 1983 AGTGGAAATCATGCTTTGGAGGAGTTGGCGAAAGCTCTTGAAAAACGCAAAATTCAGCTG 2042
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
Db 2043 GTTCTTACCAATCCCGGGCGCGGTGATCCAGAGCTCGCTCAGGAAATTCACGGAC 2102
Qy 641 HisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPhe 660
Db 2103 ATGATGGTGAAGCAACATATTTCTCAGCGTGGCGAGCGCTGTG----- 2147
Qy 661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsn 678
Db 2148 -----ANGAAATTTGCTCTTAAGATGGTCGATAAATGCTGATGGAGT 2189

RESULT 6 US-10-425-115-182207
US-10-425-115-182207, Application US/10425115
; Publication No. US2004021427A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 182207
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_97750C.1
US-10-425-115-182207

Alignment Scores:
Pred. No.: 1-5e-183 Length: 2651
Score: 1845.50 Matches: 355
Percent Similarity: 74.41% Conservative: 116
Best Local Similarity: 56.08% Mismatches: 157
Query Match: 53.11% Indels: 5
DB: 20 Gaps: 3

US-10-762-049-18 (1-680) x US-10-425-115-182207 (1-2651)
Qy 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
```

```
Db 518 GTGSCAAAGGTGGCGCGCGCGCGCGAGCACGGCGAGCAAGATGAAGGTGAGGGTG 577
Qy 61 LysGluThrPhePheProAspPheProLeuArgGlnPheLysAsnLysProAlaSerLys 80
Db 578 AAGGAGACCTTCTTCCCGCAGCAGCCGTTCCGGCGGTTCAAGGGCAGCGCCGGGAGCG 637
Qy 81 LysPheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThr 100
Db 638 CAGTGGCTCATGGCGGTAGTACCTCTTCCCATCTGAGTGGGTGGCGAGTACTCC 697
Qy 101 PheGlnPheLeuLysAlaAspLeuileAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
Db 698 TTGTCGCTCTTCAAGTCCAGCTCGTCCGCGGCTCACCATTGCCAGCTTCGCATTCCT 757
Qy 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSer 140
Db 758 CAGGCGATTAGTACGCGAAGCTGGCAAGCTTGGCTCCCATTAATCGGCTGATTTCGAGC 817
Qy 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
Db 818 TTCTGTCGCCGATGCTGACGCTGCTGCGGAGCTCCCGTACCTGGCGGTGGCGCCG 877
Qy 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
Db 878 GTGTGATCTCTGCTGATCATGGGTCCATGCTGCGGAGCGCTGAGCCCCACCGCG 937
Qy 181 AspProLysLeuTyrIleHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
Db 938 GAGCCGACGCTGTTCTCTGAGCTGGCTTCCCTCCACCTGTTCCGCGGGCTGGTGAG 997
Qy 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
Db 998 GCCTCCCTGGGATCCCTCAGGCTCGGCTTCGTCATCGACTTCTGTCCTCAAGGCGAGCTG 1057
Qy 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeu 240
Db 1058 GTGGGGTTCATGGCGCGCGCCCATCATCTGCGGCTCGAGCAGCTCAAGGCGGTGCTG 1117
Qy 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
Db 1118 GGCATCGTCCACTTCACCAACCGAGATGGGATCGTCCAGTCATGGCTCCGTTCCAC 1177
Qy 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
Db 1178 CACACGAGCGAGTGGTGGCGAGCATCTCATGGGCTCTGCTTCTCTGCTTCTGCTG 1237
Qy 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
Db 1238 CTGTGGCGAGGATGTGAGCATCAGATGCGCCAAAGCTTTTCTGGGTTCGCGGTGCGCG 1297
Qy 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db 1298 CCCTGGCATCGTCCACCATCTCGAGCTGCTGTTTCTCTTCTTCAAGCTCAGAACCAT 1357
Qy 321 GlyValGluValIleGlyLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
Db 1358 GGCATCAGCATCATTTGGGAGCTCAAGTGGCGCTGAATCGCCCTCGTGGGAGCAAGCTC 1417
Qy 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
Db 1418 CTGTTTGACACGGCGTATTTAGGCTTCACCATGAGAGCTGGGCTTGTACCCGGAATCATC 1477
Qy 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
Db 1478 TCATGACCGAAGGAATAGCGTTGGTGAACATTTGCTCTCACTCAAGGACTACAGATA 1537
Qy 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
Db 1538 GATGGAACACAGGAGATGATGCGCATAGGTTGATGAATGTTGTTGGTCTCTGCACATCA 1597
Qy 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
```

```
Db 1598 TGCTACCTAACACAGGTGGTTCCTCCGGCTCTGCTGTAAACACACAGCGCGCTCGAAG 1657
Qy 421 ThrAlaAlaSerAsnIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db 1658 ACTGCCATGTCACACGTCATCGCGCTGCTGATGGTCAGGCTGCTGCTCCTCATG 1717
Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db 1718 CCACCTGTTCTGTATACACACCAACGTTGCTCTCGGACGATCATCGCGCGGTGATC 1777
Qy 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db 1778 GGCTGATCATGATTCCCGCGGTGTACCATCTCGAAGATGACAGATGGATTTCTG 1837
Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db 1838 GTGTGGTGTGGCGGTTCGCGCGTTCATCTTCAGTCCAAAGAGCCCTTGGGATA 1897
Qy 501 AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal 520
Db 1898 GCGGTGTGTATATATATTAGGTGTGTGATGCAGATCAGGAGCCCAAGATGATGGTT 1957
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db 1958 CAAGGGAACATCAAGGGGACTGATATTTACAGAGCTCGCATCACTACACAGGAGGCCAA 2017
Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db 2018 AGAGTTCTGGGTTCCTGATCTTGGCCATTGAAGCCAGGATTAACCTTCGCCAATCCAAC 2077
Qy 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
Db 2078 TACCTGAATCAAGAGGATTAAGAGATGGATA-----GAGGAAGAATCTTTTGAACAGGAT 2131
Qy 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db 2132 AAACATACAGAACTCCATTTCCTGATTTGGATCTGTGAGCTGTCTCTGCAATTGACACA 2191
Qy 601 SerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeu 620
Db 2192 AGTGGCATAGCGTTCCTCATTTGACATPAAAGAAATCAATAGAGAAACGTGGTCTGGAGCTT 2251
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLys---SerLysPheGln 639
Db 2252 GTGCTTGTCAATCCAACTGAGGAAGTCATGGAGAAAATACACGTCGCAACAGCGCTGAA 2311
Qy 640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 659
Db 2312 AACTATTTTAGCCAGATTGCTTGTATCTGACCACCTGGCGAAGCAATC-----GCTTCA 2365
Qy 660 PheAsnLeuArgAlaSerLysThrAsnProLysLysAsp 672
Db 2366 CTTTCTGCATTGCCAAGATGACAAAACCCCTAAATGGAT 2404

RESULT 7
US-10-424-599-65103
; Sequence 65103, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 28
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 65103
; LENGTH: 2366
; TYPE: DNA
; ORGANISM: Glycine max
```

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_29800C.1

US-10-424-599-65103

Alignment Scores:

Pred. No.:	1,59e-183	Length:	2366
Score:	1844.50	Matches:	367
Percent Similarity:	73.64%	Conservative:	133
Best Local Similarity:	54.05%	Mismatches:	165
Query Match:	53.08%	Indels:	15
DB:	18	Gaps:	5

US-10-762-049-18 (1-680) x US-10-424-599-65103 (1-2366)

Qy 8 SerTyrIleThrHisIleCysLeuLeuArgAsnThrIleIleGluAspMetGlySerVal 27

Db 73 ACTTACATCACACCATTTCTCTTATA-AGTTCGTGTTCTCTTCAAAATGAGACTAAT 131

Qy 28 Asp-----TyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43

Db 132 AATGCATGTACTATGCAATTCATTCATGATAGATGTCC-----ATGGAAGTGCAACAA 185

Qy 44 ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGluThr 63

Db 186 GTTGTGCCACCACTCACAAGAGCACCTTCAGAACCTCAAGGGTAGGCTCAAGGAAACT 245

Qy 64 PhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPheMet 83

Db 246 TTCCTCCCTGATGATCTCTGCGCCATTCAGGACACCACTCTTAAGAGAAAACGTGATC 305

Qy 84 LeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPheGlnPhe 103

Db 306 CTTGGAGCTCAATATGTGTCTTCTTCCAAATGGGGCTTAAGTATATAATCTCAAACTC 365

Qy 104 LeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIle 123

Db 366 TTCAAATCTGACCTTGTCTGCGCTCACTATTGTAGCTTGGACATCCCGCAGGGAATG 425

Qy 124 SerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIlePro 143

Db 426 AGTTATGCTTAGCTTGGCAAGTCTTCTCCAAATTTAGGACTTTATTCTAGTTTGTCCA 485

Qy 144 ProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaVal 163

Db 486 CCACCTGTCTATGCTGTCTTGGAAAGCTCAAAAGGACCTTGACAGTTGGACCTGTTCTATT 545

Qy 164 GlySerLeuLeuMetCysSerMetLeuSerAsnAlaValAspProAsnGluAspProLys 183

Db 546 GCTTCTTGTGATGGGATCCATGTTGCATCAGGAAGTGTCTCCCAACACAGATCCAAAT 605

Qy 184 LeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeu 203

Db 606 CTGTTTCTTCAGCTAGCTTTCATTCACATTTATTGGCGGTCTCTTTCAAGCTTGTCTT 665

Qy 204 GlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPhe 223

Db 666 GGAATCCTAAGGCTAGGCTTCATAATTTGATTTCTATCTAAGGCCATTCTTATTGGGTT 725

Qy 224 MetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGlu 243

Db 726 ATGGCTGGAGCTGCTATTATTGCTCCTGCAACAGCTCAAGAGCTGCTTGGATTCACA 785

Qy 244 HisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHis 263

Db 786 CATTTCATTAATCAGATGGGTCTGATTCTGTATTGACTTCTGTTTTTTCACAAATATACAT 845

Qy 264 GluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThr 283

Db 846 GAGTGTGATGCAACCAATATTGATGGGATTTGCTTCTTCTGTTACTACTACTATTAGCA 905

Qy 284 ArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeuThr 303

Db 906 AGACACGTTAGCATTAAGGAAACCAAACTATTCTGGGTCTCAGCTGGAGCTCCTCTTATG 965



```
QY 304 SerValIleLeuGlySerIleuValTyrPheThrHisAlaGluLysHisGlyValGlu 323
Db 966 TGTGTCATCATCTCTACCCCTCTTGGTTTGGCAATTAAAGGCTCAAAATCATGTCATCAGT 1025
QY 324 ValIleGlyLeuLysGlyLysGlyLeuAsnProSerLeuThrAsnLeuValPheVal 343
Db 1026 GCGATTGGAAATTTGCAACAGGAATAAATCTCCATCATGGAATATGTTGCTCTTTCAT 1085
QY 344 SerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAla 363
Db 1086 GGAAGTCACCTAGGCGCTAGTTATGAAACCAAGGCGTTATCACCGGATTTTGTCCCTAAACA 1145
QY 364 -GluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAs 383
Db 1146 GGAAGGTAATTGCGAGTAGAAGGACATTTGCGAGCTCTCAAAACTCAACAAAGTGGATGAAA 1205
QY 383 nLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrIle 403
Db 1206 TAAGGAATGATGCGCAATGGGTTTATGAATGTTGTTGGCTCTTCACTTCCGTCTATGT 1265
QY 403 uThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAl 423
Db 1266 TACAACAGGTGCTTCTCGGTGAGCAGTTAAACAATGCGAGGCGCAAAAACAGCTGT 1325
QY 423 aSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheIleuThrProLeuph 443
Db 1326 GTCAAAATGATGATGCTGTGACAGTCATGGTGACACTCTTTTCTCATGTCATGTGT 1385
QY 443 eHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIle 463
Db 1386 TCATATACAGCCCTAATGTCGTGTGGCGCAATCATAGTCACAGCAGTAATGCGCTCAT 1445
QY 463 eAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMe 483
Db 1446 CGATCTCCCGCGTGTGTAAACATTTGGAAGATCGCAAAATTCGATTTTGTGTGATGAT 1505
QY 483 tSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleVa 503
Db 1506 GACTGCTTCTTAGGTGTTCTTTTCATCTCTGTCCAAAGGAGGCTTGTCTCTGCTGTGG 1565
QY 503 lIleSerValLeuArgValLeuLeuPheIleAlaArgProArgTyrPheValLeuGlyAs 523
Db 1566 ATTATCACTTTGAGATACCTCTTGCAATATAGACAGCGCAAAACAGTATGTTGGGAA 1625
QY 523 nIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValPr 543
Db 1626 GATACCAGGAACAGACATATAGAAATCTTGATCAATACAAGGAAGCTGTGAGAATACC 1685
QY 543 oGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuAr 563
Db 1686 TGGATTTCTTATTTTAAGCATTCAGGCTCCCATCAATTTTGCACCAATCATCATATCTCAA 1745
QY 563 gGluArgIleThrArgTTrpIleAspGluGluGluArgIleLysAlaIleThrGlyGluTh 583
Db 1746 CGAGAGACGTTAAGATGGATTGGAAGAAGAAGAACACATTAAGAA-----CAATT 1799
QY 583 rSerLeuGlnTyrValIleIleAspMetSerAlaValAlGlyAsnIleAspThrSerGlyI1 603
Db 1800 AAGCCTTCGATCTTAGTATTGGAATGTGAGCTGTGAGTGTGTTGACACAAAGTGGAAAT 1859
QY 603 eSerMetLeuGluValLysLysIleThrGluArgArgGluLeuLeuValLeuVa 623
Db 1860 CTCACCTTTTCAAGGAGTTGAAACCAACTCTGGAAGAAAGAGGTGTGTGAGCTGTGTGTGT 1919
QY 623 lAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn---HisLe 642
Db 1920 CAATCTCTTCGCTGAGTCATAGAAAGCTTAAAGAACGACAGATGAAGCTAATGATTTCAAT 1979
QY 642 uGlyLysLysTrpIleTyrLeuThrValGluGluAlaValAlGlyAlaCysAsnPheAsnLe 662
Db 1980 AAGAGCAGATAACCTTTTCTTGACAGTTGGAGAGGCTGTAGCT-----TCACCT 2027
```

```
QY 662 uArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTTrpAsnAsnVal 680
Db 2028 TTCTTCAGCAATGAAGAGGCAATCATCAACCATTTACAGAAGGGGCACACACAATT 2082

RESULT 8
US-10-437-963-95853
; Sequence 95853, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 95853
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94005C.1
US-10-437-963-95853
```

```
Alignment Scores:
Pred. No.: 3,69e-180 Length: 2001
Score: 1811.50 Matches: 340
Percent Similarity: 73.10% Conservative: 122
Best Local Similarity: 53.80% Mismatches: 167
Query Match: 52.13% Indels: 3
DB: 19 Gaps: 3
```

US-10-762-049-18 (1-680) x US-10-437-963-95853 (1-2001)

```
QY 43 GlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db 106 GAGGTGAACCTGTCCGGGCGCGCCGTTCCGGGAGAACCTGTGTCGGACCTCGCGGAG 165
QY 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db 166 ACGTTCTTCCCGACGACCGGTTCCGGCGGTTCCGGCGCTGCCCGCGCGCGCGCGCG 225
QY 83 MetLeuGlyLeuGlnPhePhePheProIlePheGluTTrpAlaProLysTyrThrPheGln 102
Db 226 TGGTGGCGGTCAAGTACTTTCGCGCGCTGGACTGGGTCCCGGCTACGSCCTCGAC 285
QY 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 286 AAGTTCAGATTTCGACCTCTCTCCGGCATCACATGCCAGCCTCGCCATCCGCGAGGC 345
QY 123 IleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIle 142
Db 346 ATCAGCTACGCGCGCTCGCAACCTCCCTCCAATCATCGGCTCTATTTCGAGCTTCGTG 405
QY 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 406 CGCGCGTTCATGTACGCGGTTCGGGAGCTCCAAACACCTGCGCGTGGGACGCGTGGCG 465
QY 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
Db 466 GCGGCGTCTGCTGCTGGCGTCCATCATCGACGAGGCTGGCGCGGACGAGACCCG 525
QY 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 526 CAGCTGTACTTCAGCTCTTCTTACACCGCGCGCTTCTTACCGGCTCTTCCAGACCGCG 595
```



```
QY 203 LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleLeuGly 222
Db CTAGGCGTCTTTAGGTTAGGCTAATAGTGAATTTCTCTCGGGTCGACGATCACCGG 645
QY 223 PheMetGlyClyAlaIleThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeu 242
Db TTTCATGGCGGCACCGCGATGATCATACTACAGCAGTTCAAGGCGCTCTCTGGGATG 705
QY 243 GluHisPheThrHisGlyAlaAspIleLeuSerValMetArgSerValPheThrGlnThr 262
Db AAGCACTTACCAACCAAGACAGACATCATCTCGTCTCGACCTACCACTACCATACCGG 765
QY 263 HisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSer 282
Db CACGAGTGAATGGCAGAGTGCGAGTTCTCGGCATATGCTTCTCTGTTCTTGATGTCG 825
QY 283 ThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeu 302
Db AGTAAGCACTCGAAGAAAGAACTGCCAAAGTTGTTCTGGGTGTCGGCCATTGGCCATTC 885
QY 303 ThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVal 322
Db ATGGTTGTGTCATCGAGGCATCTTCGCTTTCCTTGTCAAGGTGACGAGCATGGAAT 945
QY 323 GluValIleGlyLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPhe 342
Db CCGATAGTCGGTAACCTGAATAAGGATCAATCCCTGTCCATTGCGCAGCTTAAGTTTC 1005
QY 343 ValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleLeuSerLeu 362
Db CAGTCCAGGCACATGAAGACGGCAGTGAAGCTGGTCTCTGTCTGGGATCTCTAGCACTG 1065
QY 363 AlaGluGlyIleAlaValAlaGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGly 382
Db GCAGAAGGAATACCGTCCGCCAAGCTTGGCGATGGTCAAGAACGAGCAGATCGACGGC 1125
QY 383 AsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr 402
Db AACCAAGGAGATGATCGGTTGGTATCATGAACATCGCGGTTCTTTCACCTCTGCTAC 1185
QY 403 LeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAla 422
Db CTCACAGACAGGCGGTTCTCAAAGTCGGCGGTGAACCTTCACGCCGGGTGCAAGACGCG 1245
QY 423 AlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeu 442
Db ATGTCGAAGCTGGTGATGCGGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1305
QY 443 PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu 462
Db TTCAAGTACACCCCGCTGGTGGCGCTCTCTCCATCATGTCGTGCCATGATCGGGCTC 1365
QY 463 IleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCys 482
Db GTCAAGTCAAGGAGTTCGCGCACCTCTCAAGGTGCAAGTTCGACATCTGTCATCTGC 1425
QY 483 MetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle 502
Db ATGGTCGCTTCTCGCGCTGCTGCTTCTTCACCATGTCGAGGCTTCAGCGCATCCGTA 1485
QY 503 ValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuGly 522
Db GGCTTGTCAAGTTGTCAGGGGCTTGTGTAGTGGCCAGGCTGCCACCTGCAAGCTCGGG 1545
QY 523 AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVal 542
Db AACATAGCAGGAGGAGAGACCTTCCCGCAGGTGAAGCAATATACCCCAAGCAAGAGCATC 1605
QY 543 ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu 562
Db CTGGGATCTCTGCTTGGAGCTGCGCTCTCCCATCTACTTCTGTCATGCGGGTTACCTG 1665
QY 563 ArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGlu 582
```

```
Db 1666 CGAGAAAGGATTTTGAGATGGGTGAAGATGAGGATAACCTGTGCAAGAGCGTCGGGCAC 1725
QY 583 ThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly 602
Db ---GATCTGCAATACTGTGTTCTTGTATCTGGTGTGTCACCTTCTGTGCACAACTCTGGA 1782
QY 603 IleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeu 622
Db GTTGGGATGCTACTAGAAAGTACACAGAGCCCTCGAAGAGAGGATACGATAGTCTGTG 1842
QY 623 ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641
Db ACGAATCCGAGGCTAGAGGTAAACAGAGAAGCTGGTGTCTCTGGATACGTACAGGACATC 1902
QY 642 LeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661
Db TTAGGGGATGAGTGGTCTTCTGACGGTCAAGGACGCCATCACGGCGGTGTCGATACGGC 1962
QY 662 LeuArgAlaSerLysThrAsnProLysLysAspGlu 673
Db CTGCGAGATATCCAGA--ATAAGGGAGAGACGAA 1995
RESULT 9
US-10-437-963-77005
; Sequence 77005, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77005
; LENGTH: 2780
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MFT4530_76946C.1
US-10-437-963-77005
Alignment Scores:
Pred. No.: 6,49e-176 Length: 2780
Score: 1773.50 Matches: 328
Percent Similarity: 73.13% Conservative: 128
Best Local Similarity: 52.65% Mismatches: 166
Query Match: 51.04% Indels: 1
DB: 19 Gaps: 1
US-10-762-049-18 (1-680) x US-10-437-963-77005 (1-2780)
QY 34 GlyMetAsnAsnPheGluArgValHisGlnValGluValProProGlnProPhePhe 53
Db GGCCATCATCATCATCATCACGGCCACAAAGTTCGAGTTCCCAACCAAGAGAGTCTCATC 492
QY 54 LysSerLeuLysTyrSerLeuLysGluThrPhePheProAspAspProLeuArgGlnPhe 73
Db GACGAGTTTCACCGACCGCGTGAAGAGAGAGCTTCTTCGCGACGACCGCTTGGGAGTAC 552
QY 74 LysAsnLysProAlaSerLysLysPheMetLeuGlyLeuGlnPhePheProIlePhe 93
Db AAGGACCAAGCCCATGTCATAAGAGAGGTGTGTATCACCCTGCAAACTTCTTCCCGGTGCTG 612
QY 94 GluTrpAlaProLysTyrThrPheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThr 113
```

Db	613	GA	CTGGGGCGG	CACTAC	CTCCG	CAAGTT	CTAGGG	CGCACT	CTGCT	CTCGGC	CTCACC	672	
Qy	114	Ile	AlaSer	LeuAla	IlePro	GlnGly	IleSer	TyrAla	ValSer	LeuAla	AsnLeu	ProPro	133
Db	673	AT	TGCCAG	CTCT	GCAT	ACCC	AGCAT	CGGT	TATG	CGAAG	CTTG	TGCTGC	732
Qy	134	Ile	LeuGly	LeuTyr	SerSer	PheIle	ProPro	LeuIle	TyrAla	MetMet	GlySer	Ser	153
Db	733	AA	TATGG	ACTGT	ACAG	CAGCTT	CGTAC	CGCCTT	TGAT	TATAC	CGCAT	GTATGG	792
Qy	154	Arg	Asp	LeuAla	ValGly	ThrVal	AlaVal	GlySer	LeuLeu	MetGly	SerMet	LeuSer	173
Db	793	AG	GATAT	AGCC	ATTG	TCCAG	TGGCG	TGTT	CGCT	GCTG	CTGCA	CTCTCT	852
Qy	174	Asn	AlaVal	AspPro	AsnGlu	AspPro	LeuLys	LeuTyr	LeuHis	LeuAla	PheThr	AlaThr	193
Db	853	AA	TGAGT	TTT	GATCC	CAAG	AGAAT	CAG	AAG	AGTAC	CCCG	TGGCTT	912
Qy	194	Leu	PheAla	GlyVal	PheGln	AlaAla	LeuGly	LeuPhe	ArgLeu	GlyLeu	IleVal	Asp	213
Db	913	TT	CTTC	GCTG	GGGT	CACC	AGCAG	TGCT	CGG	ATTCT	CAGG	CTAGG	972
Qy	214	Phe	LeuSer	HisAla	ThrIle	IleGly	PheMet	GlyGly	AlaAla	ThrVal	ValCys	Leu	233
Db	973	TT	CTTGT	GCC	ATGCT	GCC	ATGTT	GG	AA	TTCAT	GCGG	AGCG	1032
Qy	234	Gln	Gln	LeuLys	SerIle	LeuGly	LeuGlu	HisPhe	ThrHis	GlyAla	AspIle	Ser	253
Db	1033	CAG	CAGCTT	TA	AGCCTT	CTCT	TGG	AA	TTCAC	CAAG	AGACT	GATCAT	1092
Qy	254	Val	Met	Arg	SerVal	PheThr	GlnThr	HisGlu	---Trp	ArgTrp	GluSer	AlaVal	272
Db	1093	GT	CATG	AAAT	CAGT	CTCG	GGAA	TGTT	CAC	CATGG	TGGA	CTGG	1152
Qy	273	Gly	CysVal	PheIle	PhePhe	LeuLeu	SerThr	ArgTyr	PheSer	LysLys	ArgPro	Arg	292
Db	1153	GG	AGCAT	CTCT	CTCG	GG	ATTCT	CTCT	CGT	TGCA	AAATAC	ATTG	1212
Qy	293	Phe	PheTrp	ValSer	AlaMet	AlaPro	LeuThr	SerVal	IleLeu	GlySer	LeuLeu	Val	312
Db	1213	CT	CTTT	TGG	TGGT	CGCA	ATTG	CAC	CACT	CTCT	CGGT	ATCAT	1272
Qy	313	Tyr	PheThr	HisAla	GluLys	HisGly	ValGlu	ValIle	GlyGlu	LeuLys	GlyLys	GlyLeu	332
Db	1273	TAT	ATCA	CTCG	AGCAG	CAAG	CTGGT	GTCT	GTG	ATTG	TCTCA	AGTAC	1332
Qy	333	Asn	ProPro	SerLeu	ThrAsn	LeuVal	PheVal	SerPro	TyrMet	ThrThr	AlaVal	Lys	352
Db	1333	AA	TCAC	CTT	CAG	CTAG	CTCAT	ATAT	CTT	CAG	TGG	CCCA	1392
Qy	353	Thr	GlyIle	ValVal	GlyIle	SerLeu	AlaGlu	GlyIle	AlaVal	GlyLeu	LysGly	SerPhe	372
Db	1393	AT	AGG	ATG	ATAG	CTGG	CATG	ATAG	GCGCTT	CA	AGAAC	AAAT	1452
Qy	373	Ala	MetTyr	LysAsn	TyrAsn	IleAsp	GlyAsn	LysGlu	MetIle	AlaIle	GlyThr	Met	392
Db	1453	GCT	GGCCT	GA	GGACT	CA	AGATT	GAC	GGAA	CA	AGGA	AAAT	1512
Qy	393	Asn	ValVal	GlySer	PheThr	SerCys	TyrLeu	ThrThr	GlyPro	PheSer	ArgSer	Ala	412
Db	1513	AA	CATT	TGT	TGTT	CAAT	GACT	TTCT	TGCT	AGT	TATAC	CGG	1572
Qy	413	Val	AsnTyr	AsnAla	GlyCys	LysThr	AlaAla	SerAsn	IleIle	MetSer	LeuAla	Val	432
Db	1573	GT	CAAT	TAC	ATG	TGGT	CGC	CAAC	AGCG	GTAT	CCCA	ATAT	1632
Qy	433	Met	LeuThr	LeuLeu	PheLeu	ThrPro	LeuPhe	HisTyr	ThrPro	LeuVal	ValLeu	Ser	452
Db	1633	CT	GTCT	CAC	ACT	GGAG	CTG	ATTCT	CA	ATTG	TTC	AGT	1692
Qy	453	Ala	IleIle	ValSer	AlaMet	LeuGly	LeuIle	AspTyr	GluAla	IleHis	LeuPhe	Thr	472

Db	1693	TCCATCATCATATACAGCAGTGCTTGGCTTAGTCGACTACACACAGCAATACCTTATCTCGG	1751
Qy	473	LysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValValPheGly	492
Db	1753	AAGGTTGATAAATTGACTTCTGGCATGCTTAGGAGCAATCTCTTGGCGTCATATTTC	1812
Qy	493	SerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuLeuPhe	512
Db	1813	TCAGTGAGGATGCTGCTGTAATGGCGTTGCCATCTCACTCGCTAAATATCTTCTACAA	1872
Qy	513	IleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleIleTyrArgAsn	532
Db	1873	GTTCACGACCCAGAACAGATTTTGTCTGGTAACCTTCCCAAGAACACATATATAGGAAT	1932
Qy	533	ValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAla	552
Db	1933	ATCGACCAGTATCCAGAAAGCTACTCTTGGTACCAGGAGTGGTAATAGTTAGGGTGGACTCA	1992
Qy	553	ProIleTyrPheAlaIleAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGlu	572
Db	1993	GCTATCTACTTCAAACTCGAACTATGTTAAAGACAGAATCTCTGAGGTGGCTGAGGGAC	2052
Qy	573	GluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMet	592
Db	2053	GAGGAAGACGTCCAGCAGGAACAGAGCTACAGAAACTGAGTTCTGTATGTTGAGCTA	2112
Qy	593	SerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLysIle	612
Db	2113	TCTCCGGTAATTGCATTTGATACAGTGGAACTCATGCCTTGGAGGATCTATTAGAGCA	2172
Qy	613	ThrGluArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLysLys	632
Db	2173	CTTGAAAGCGCAAAATCCAGCTGATCTTAGCCAAATCCGGGGCGAGCTGTGATCCTGAAG	2232
Qy	633	LeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIleTyrLeuThrValGlu	652
Db	2233	CTCCGATCAGCGAAATTCACGGATCTCATCGGTGAAGACAAAGATATTCCTGACTGTGGG	2292
Qy	653	GluAlaVal 655	
Db	2293	GAGCGAGTG 2301	
RESULT 10			
US-10-425-114-15719			
; Sequence 15719, Application US/10425114			
; Publication No. US20040034888A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Jingdong			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Screen, Steven E			
; APPLICANT: Tabaka, Jack E			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(53313)B			
; CURRENT APPLICATION NUMBER: US/10/425,114			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 73128			
; SEQ ID NO 15719			
; LENGTH: 2656			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; OTHER INFORMATION: Clone ID: LIB3060-059-F10_FU1			
US-10-425-114-15719			
Alignment Scores:			
Pred. No.:	9.87e-175	Length:	2656
Score:	1762.00	Matches:	334
Percent Similarity:	73.23%	Conservative:	120
Best Local Similarity:	53.87%	Mismatches:	162
Query Match:	50.71%	Indels:	4

Alignment Scores:		
Pred. No.:	9.87a-175	265b
Score:	1762.00	334
Percent Similarity:	73.2%	120
Best Local Similarity:	53.8%	162
Query Match:	50.71%	4
		Indels:
		Mismatches:
		Matches:
		Length:

DB: 18 Gaps: 2  
US-10-762-049-18 (1-680) x US-10-425-114-15719 (1-2656)  
QY 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60  
DB 299 CTGCAAGAGGTGTCGTCCGGAGCGCGACGACGCGAAGCGCTGCGGCACGCCCTG 358  
QY 61 LysGluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80  
DB 359 GCCAGAGGTTCCTCCGGAGACCCGCTGCACAGTTCAGAACACAGTCTCGCGCGCG 418  
QY 81 LysPheMetLeuGlyLeuGlnPhePheProPheProLeuTrpAlaProLysTyrThr 100  
DB 419 CGCCTCGTGTGGCGTGCACACTACTTCTCCCATCTTCAGTGGGGTCCGCGCTACAGC 478  
QY 101 PheGlnPheLysAlaAspLeuLeuAlaGlyLeuThrLeuAlaSerLeuAlaLeuPro 120  
DB 479 CCAGCGCTCTCGGCTCCGACCTCGTCGGCGGCTCACCATTGCCAGCTCGCCATCCG 538  
QY 121 GlnGlyLeuSerTyrAlaLysLeuAlaAsnLeuProProLeuGlyLeuTyrSerSer 140  
DB 539 CAGGGAATCAGTACGCGACAGCTCGCAACCTCCGCCAATCTTGGCTATATCCAGC 598  
QY 141 PheLeuProLeuLeuTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160  
DB 599 TTCGTGCGCGCTCATCTACGCGCTGCTGGGAGCTCGCGGACCTGGCGGTGGGCGC 658  
QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180  
DB 659 GTGTCCATCGCGTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 718  
QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200  
DB 719 CAGCCGCTCTTACCTTCAGCTGAGCTTCCAGCCACCTTCTCCGCGCGTCTTCCAG 778  
QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuLeuValAspPheLeuSerHisAlaThrLe 220  
DB 779 GCCTCTCCGATTCCTCAGCTGGGCTTCATCTGTCGACTTCCTGTCCAAGGCGACGCT 838  
QY 221 IleGlyPheMetGlyAlaAlaThrValValCysLeuGlnLeuLysSerLeuLeu 240  
DB 839 ACGGGCTTCATGGCGCGCGCGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 898  
QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleLeuSerValMetArgSerValPheThr 260  
DB 899 GGCACTCCCACTTACCTCCACATGGGATTCCTCGAGCTCATGCGCTCCGCTGCTCAAC 958  
QY 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheLeuPheLeu 280  
DB 959 CGCCACGACGAGTGAAGTGGCAGACGATGCTCATGGGCTCCGCTTCCTCGCCATCCTC 1018  
QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300  
DB 1019 CTCTCTACGCGCCAAATCAGCGCCAGGAACCCCAAGCTTTTCTGGGTATCAGCAGGTG 1078  
QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320  
DB 1079 CCCCTGGCGTGGGTGATCATCTCCACATCTCTCTCTCATCTGG-----AAATCCCC 1132  
QY 321 GlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340  
DB 1133 AGCATCAGTGTATTGGCATCTCTCCAGGGAGTGAACCTCTCTTCGGGAACATGCTC 1192  
QY 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyLeuValValGlyIleIle 360  
DB 1193 AGCTTCAGCGGCTCCTATGTGGCGCTGCACATCAAAACCGGGATCATGACAGCATCTG 1252  
QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380  
DB 1253 TCCTTAACAGAGGATCGCAGTGGGAGGACCTTCGCTGCCATCAACACTACAGGTG 1312  
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400

DB 1313 GACCGGAACAAGAGATGATGCGATCGGCTGATGAACATGGCGGCTCTCTGCGCTCC 1372  
QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420  
DB 1373 TGCTACGTGACACAGCGGGTCTTCTCCCGGTGCGGGGTGAATACAGCGGGCTGACAG 1432  
QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440  
DB 1433 ACGGCGCTGCTCAACGTCGTGATGGCGGCGGTGCTGGTGCAGCTGCTGTCTCTCATG 1492  
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460  
DB 1493 CGCGCTTTCACATACACCCGAACGTGATCTCTGGGCGGATCATCATCAGCGGCTGGTG 1552  
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480  
DB 1553 GGGCTGGTGGAGCTGCGCGGCGCCGAGGCTGTGGAAGTGGACAAGTGGACTTCCTG 1612  
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500  
DB 1613 GCGTGGTGGCGGCTTCTCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1672  
QY 501 AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal 520  
DB 1673 GCGCTGGCATCTCGCTCTTCAAGTCTCTGTGCAAGTCCCGCCGCCCAACGTCGTGGTG 1732  
QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540  
DB 1733 GAGGCGCTCGTCCCGGGACGACAGACTACCGCAGCGTGGCGCAGTACCGGAGGCCGTC 1792  
QY 541 HisValProGlyMetLeuLeuLeuIleAspAlaProIleTyrPheAlaAsnAlaSer 560  
DB 1793 CCGCTGGCGGCTTCTCTGCTGCGGCTGAGTCCGCGCTCTACTTCGCCAATCCATG 1852  
QY 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluGluArgIleLysAlaThr 580  
DB 1853 TACCTGGTGGAGCGGCTCATGCGCTACCTCCGACGAGGAGGCGCGCTCAAGTCC 1912  
QY 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600  
DB 1913 AACCACTCTCCATCCGATGCGTCTGCACTGGCGGCTCGCGCGATCGACAG 1972  
QY 601 SerGlyIleSerMetLeuGluGluValLysIleIleThrGluArgGluLeuGlnLeu 620  
DB 1973 ACGGCTTAGACGCGCTGTCGAGCTCAAGAAAGTCTTGACAAAAGAACATCGAGCTG 2032  
QY 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640  
DB 2033 GTGCTTGCCCAACCCGCTGGGTGGCGGAGAGATGTTTCAACTCGCGCGTGGCGGAG 2092  
QY 641 HisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaVal-----GlyAlaCys 658  
DB 2093 AGCTTGGGTGCGGCGCTCTTCTTTCAGCTAGCGGAGGCCCTCGCGCGGGGGCTGC 2152  
RESULT 11  
US-10-425-115-75763  
; Sequence 75763, Application US/10425115  
; Publication No. US2004021472A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222) B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 75763  
; LENGTH: 2713  
; TYPE: DNA



```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 73410
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73694C.1
; US-10-437-963-73410

Alignment Scores:
Pred. No.: 6.84e-172 Length: 2429
Score: 1734.50 Matches: 333
Percent Similarity: 70.93% Conservative: 116
Best Local Similarity: 52.61% Mismatches: 163
Query Match: 49.91% Indels: 21
DB: 19 Gaps: 3

US-10-762-049-18 (1-680) x US-10-437-963-73410 (1-2429)

Qy 41 ValHieGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2261 CTCGACAAAGGTGTCGGTGGCGGAGAGGAGGTGCGACGCGGAGCGCTGAGCGCGGCTG 2202

Qy 61 LysGluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2201 GCGGAGGTGTTCTTCCCGGACGACCGCGCTCCACAGTTCAAGAACACAGTCGTCGCGCGG 2142

Qy 81 LysPheMetLeuGlyLeuGlnPhePheProilePheGluTrpAlaProLysTyrThr 100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2141 CGGCTGGTCTCGGCTGCGAGTACTTCTTCCCATCTTTCATTTGGGGTCCGACTACAGC 2082

Qy 101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2081 CTCGCGCTCTCGCTCCGAC----- 2061

Qy 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSer 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2060 ---GGAATCAGCTACGCGCAAGCTCGCAACCTCGCTCCAATCATTTGAGCTATATTCGAGC 2004

Qy 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2003 TTGTCGCGCGCTGATCTACTCTGTTGGTAGCTCGCGGATCTGGCTGTAGGGCGG 1944

Qy 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValaAspProAsnGlu 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1943 GTGTCGATACGTCGCTGGTGATGGGTGATGTCGCGGAGCGGTGTCGCGGACCAG 1884

Qy 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1883 GAGCCCATCTCTACCTCCAGCTCGCCTTCCACCTCCACCTTCTTCGCGCGCTTCCTCAA 1824

Qy 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValaAspPheLeuSerHisAlaThrIle 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1823 GCCTCCCTCGGCTTCTCCGGCTGGGGTTTCATCTGGATTCTTGTGGAAGGCGAGCGCTG 1764

Qy 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1763 ACGGGAATTATGCGCGCGCGGATCATCGTGTGCTGCGAGCAGCTCAAGGGATTGCTC 1704

Qy 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1703 GGGATCATTCATTTCACGTCGCGAGATGGGGTTCGTCGAGGTGATGCACCTCCGCTTTCAG 1644
```

```
Qy 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1643 CACCACGAGTGGCGTGGCAGACCATCTCTATGGGGGTCCCTTCCTCCCGCTCTC 1584

Qy 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1583 CTCACCACAGCCCATCATCAGCGCCAGAACCCAAAGCTTTTCTGGGTGTGACGAGTGCT 1524

Qy 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1523 CCACCTGACATCAGTGCATCTCTACCATCTCTCGTTCGTGAGCAAGCT-----CAT 1470

Qy 321 GlyValGluValIleGlyLeuLysGlyLeuAsnProPheSerLeuThrAsnLeu 340
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1469 GGCATCAGTGTATTGGCGATCTCCCAAGGATTTGAACCTCTTCACGGAATATGCTG 1410

Qy 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleLeu 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1409 ACCTTCAGTGGCTCTACGTAGGACTGGCTTTAAACACAGGAATTATGACTGGCATACTA 1350

Qy 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1349 TCTCTCATTGAAGGATAGCAGTAGCAGACATTTGCATCCATCAACAACATACCAGGTT 1290

Qy 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValIcLysSerPheThrSer 400
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1289 GATGGAAACAGGAGATGATGGCCATTGGTGTCTGATGAACATGGCTGGCTCATGCTCCTC 1230

Qy 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1229 TGTATGTCCACACAGGATCGTTCGAGTTCGAGTGGCGGTGAACTACAGCGCGGGTGAAG 1170

Qy 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1169 ACGGCGGTGTCAACATCGTATGGCTGGCGGTGCTGGTACGCTGTGTTCCTGATG 1110

Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1109 CGGCTGTTCCATACACCCCAACGTGATCTCTCGCGATCATCATCACCGCGGTGATC 1050

Qy 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1049 GGCCTCATCGAGTCGCGGCGCGCCAGCGGTGTGGAAGTGCAGAACGTCGACTTCCTC 990

Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 989 GCTGTCATGGCGCTTCCTCGCGCTCTCTCGTCTCGTCCAGATGGGCTTCGCCATC 930

Qy 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 929 GCGTCGCGATCTCTCTTCAAGATCTCTCCAGGTACCCGCGCCCAACATGGTGGTGC 870

Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 869 AAGGGCGTCTCCCGCACGCGAGTACCAGCATCGCGAGTACGCGAGGAGGCGCATG 810

Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 809 CGTGTGCGCTCTCTCGTTCGCGGTGAGTCCGCATCTACTTCGCCAATCTCATG 750

Qy 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 749 TACCTCGCGAGAGGATCATGAGGTTCTCCCGGAGGAGAGACAGCGGCGCCCAAGTGC 690

Qy 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 689 AACAGTGCCTTGTGAGATGCATCTCTCGACATGAGTGTGTGTCGACGCGATCGACAG 630

Qy 601 SerGlyIleSerMetLeuGluGluValLysIleThrGluArgGluLeuGlnLeu 620
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 629 AGTGGCTCGATGCATTCGAGAGTGTGAAGAGGTGCTGGAAAAAGAAACATCGAGCTT 570
```





```
Db 1936 TTTACTTGGCAACCTTCCAGAAACGATATCTATAGAAATGTTGAACAATATCCAGAAGC 1995
Qy 539 aLysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAl 559
Db 1996 TACCAAGTGCCAGGAGATGCTAATTGTAGAGTGGACTCAGCAATATATCTTCACAAACTC 2055
Qy 559 aSerTyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAl 579
Db 2056 CAATTATGTTAAAGAAAGATGCTGAGATGGCTGAGAGATGAGGAAGACATCAAAAGGA 2115
Qy 579 aThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAs 599
Db 2116 ACAGAAGTTACCAAAAATTGAGTTTCTGATTGTGTGACCTATCTCTGTAATGATATTGA 2175
Qy 599 pThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgGluLeuGlu 619
Db 2176 CACAAGTGGAAATCCATGCCCTTCAAGAGATTGTTGAGGACACTTGAAGAGCGCCAGATTCA 2235
Qy 619 nLeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlu 639
Db 2236 GCTGATTTTCCCAATCCTGGGGGGCTGTGATCCAAAGCTCCGGTCAGCAAAATTCAC 2295
Qy 639 nAsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluAlaVal 655
Db 2296 AGAGCTCATTTGGTGAAGAAAGATATGCTCTGACAGTTGGTGACGCCGTA 2344
```

## RESULT 14

```
US-10-424-599-131117
; Sequence 131117, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131117
; LENGTH: 1518
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89405C.1
US-10-424-599-131117
```

```
Alignment Scores:
Pred. No.: 4,22e-166 Length: 1518
Score: 1676.50 Matches: 344
Percent Similarity: 85.82% Conservative: 19
Best Local Similarity: 81.32% Mismatches: 28
Query Match: 48.24% Indels: 34
DB: 18 Gaps: 2
```

US-10-762-049-18 (1-680) x US-10-424-599-131117 (1-1518)

```
Qy 258 ValPheThrGlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIle 277
Db 51 GTTTTTTCCAAACGATGATGAGTGGAGTGGGAAAGCGCTGTGTGGTGTGTGCTTTAAT 110
Qy 278 PhePheIleuLeuSerThrArgTyrPheSerLysLysArgProArgPheThrValSer 297
Db 111 TTCTTCTCTCGTCACAGATACCTCAGCAAAACGACAGCAAAAGTTCTTGGGGGTCA 170
Qy 298 AlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAla 317
Db 171 GCATGGCCACCATGACGCTCCGTATATATGGGAAGTCTGCTGCTCTATGTGCACACAGCT 230
Qy 318 GluLysHisGlyValGluValIleGlyGluLeuLysGlyLeuAsnProProSerLeu 337
```

```
Db 231 CAGAAACATGGAGTTCAAGTGTAGTAAATTTGAAGAAAGGGGCTGAATCCACTTCGGCA 290
Qy 338 ThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValVal 357
Db 291 ACGGATTTTGGTGTGTGTGTACCTTACATGGGCACTGCCTATAAAACCTGGATTCGTACT 350
Qy 358 GlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsn 377
Db 351 GGCATTATAGCACTTGGCGAGGATAGCAGTGGGAGAGAGCTTTTGAATGTTTAAAAAT 410
Qy 378 TyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySer 397
Db 411 TATCATATTGATGGCAACAAAGAGATGATAGCTATTGGAAACCATGAATATTTTCGATCG 470
Qy 398 PheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAla 417
Db 471 TTCCTTCTTGTGTATCTCACCACCTGGACATTTTCGGGTGAGTGAACATACAAATGCT 530
Qy 418 GlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeu 437
Db 531 GGGTGCAGACAGCTGCATCCAAACATCATATGGCAATTGCAGTAATGTTGACATTGTTA 590
Qy 438 PheLeuThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSer 457
Db 591 TTCTTAACACCAATTGTTCCATTACACTCCCTCGTGTGTGCTATACAGCCATTATTGTATCT 650
Qy 458 AlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPhe 477
Db 651 GCAATGCTTGGCCTCATAGATTATAGAGAGCCATCCATCTATTTAAGGTGCACAAATTT 710
Qy 478 AspPheValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGly 497
Db 711 GACTTTGTCGTGTCAGTGCATACGTCGCGTGGTGTCTTTGGCAGTGTGAAATTGGC 770
Qy 498 LeuValIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArg 517
Db 771 TTAGTCATAGCTATTGGCATATCTGACTTCGGGTACTTCTATTATTATTCGAAGGCGAG -- 828
Qy 518 ThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGln 537
Db 829 ACATTTGTTTTGGGAACATCCAAATTTCTGTGATATACAGAAATGTTGAGCAGCTATCCA 888
Qy 538 AsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAla 557
Db 889 AATGCAAAACATGTTCTTGGATGCTAATTTCTAGAGATTGATGCACCAATTTTACTTTGCC 948
Qy 558 AsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIle 577
Db 949 AATGCAAGCTATTTAAGAGAAAGGATCA -- CGATGGATTGATGAAGAGAGAAAGAATC 1006
Qy 578 LysAlaThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsn 597
Db 1007 AAAGCTT-----GGTATG 1018
Qy 598 IleAspThrSerGly-IleSerMetLeuGluGluValLysLysIleThrGluArgArgI 617
Db 1019 TTTGATATGAGTGTGACGTGTATGCTTGGGAGATT----- 1055
Qy 617 uLeuGlnLeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerIly 637
Db 1056 -----CCAAAGAGCGAGGTGATGAAGAGAGCTGAACAAATCCAA 1093
Qy 637 sPheGlnAsnHisLeuGlyLysTrpIleTyrLeuThrValGluGluAlaValGlyAl 657
Db 1094 GTTCTTGTGAACCTGGGGCAAAATGGATCTATCTGACAGTTGAAGAGGCTGTGGTGC 1153
Qy 657 aCysAsnPheAsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTr 677
Db 1154 GTGCAACTTCATGCTCCATCTCTACAAACCGAACCCCATGAAGGATGAATCAGAAGGTTG 1213
Qy 677 pAsnAsn 679
```





Qy	580	ThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAsp	599
Db	1696	GTCCCGCGCAGGACCTGCTAGTCTGCTCGACATCGCGCGCTGACGCGCATCGAC	1755
Qy	600	ThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGln	619
Db	1756	AGCCCGGATCGAGATGCTGCGGAGGTGCACCGCGAGCTGGAGAGGAAGGGGATGAAG	1815
Qy	620	LeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGln	639
Db	1816	ATGCGGTGACGACCCGAGGATGCGGTGGCGGAGAAGCTGGTGTCTCCGGGCTCGCG	1875
Qy	640	AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn	659
Db	1876	GAGCTCTCGCGCAGAGCTGGATGTTCTCTCCAATGGCGACGCCTTGGCCCGGTGCCCG	1935
Qy	660	PheAsnLeuArgAlaSerLys	666
Db	1936	TACACGCTCCAGGGCTCCAAG	1956

Search completed: August 31, 2005, 16:25:53  
Job time : 1182 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 31, 2005, 11:52:07 ; Search time 5789 Seconds

(without alignments)  
4471.187 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARLSYTHICLLRNTI.....NLRASKTNPKKDETEGNNV 680

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10762049/runat\_30082005\_121245\_5984/app.query.fasta\_1.839  
-DB=EST -OPMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10762049 @CGN 1.1 3437 @runat\_30082005\_121245\_5984 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hrc.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_ges1.\*  
9: gb\_ges2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	2416	69.5	CNS0A7F1	BX822461 Arabidops
2	1845.5	53.1	AV105934	AY105934 Zea mays
3	1813.5	52.2	CL959635	CL959635 OsIFCC003
4	1750	50.4	CL959635	BX832405 Arabidops
5	1720.5	49.5	CL959635	CL959635 Arabidops
6	1698	48.9	CL959635	CL959635 Arabidops
7	1474	42.4	CNS0A1L2	CL961102 OsIFCC005
8	1301	37.4	CB892638	BX832587 Arabidops
9	1223.5	35.2	CK269206	CK892638 EST645430
				CK269206 EST15284

10	1135	32.7	816	6	CB624151	CB624151 OSIRIa12G
11	1122	32.3	828	7	CV130479	CV130479 X3P02a03
12	1078	31.0	831	7	CO107436	CO107436 GR_EB003
13	1016	29.2	706	7	CK102483	CK102483 GR_EB002
14	1013	29.2	783	7	CK256297	CK256297 EST739934
15	1000	28.8	785	7	CF438329	CF438329 EST674674
16	994	28.6	768	7	CN190878	CN190878 UCRCS06_0
17	983	28.3	993	6	CA133402	CA133402 SCJFRT100
18	970	27.9	583	2	AW598447	AW598447 EJ91904.Y
19	961.5	27.7	749	1	AJ794089	AJ794089 AJ794089
20	955.5	27.5	1534	5	BU693126	BU693126 QY0156_P
21	942	27.1	761	5	BQ851246	BQ851246 QY014M22_P
22	930	26.8	891	6	CD441698	CD441698 EL01N0561
23	929.5	26.7	985	9	CG279198	CG279198 OG0DS74TV
24	924	26.6	738	5	BU025764	BU025764 QHGL1F06
25	922	26.5	682	6	CA484924	CA484924 WHE4312_C
26	917	26.4	1043	7	CK163232	CK163232 FGAS01585
27	916	26.4	594	7	CO982632	CO982632 GMB9016A2
28	916	26.4	673	7	CV474471	CV474471 22850_1_D
29	915	26.3	704	6	CA220343	CA220343 SCSGFL403
30	911	26.2	728	5	BU025129	BU025129 QHF7N01.Y
31	910	26.2	745	5	BQ996962	BQ996962 QG614E18
32	908	26.1	652	6	CB035946	CB035946 VVA015H09
33	904	26.0	728	2	AW036000	AW036000 EST282859
34	902.5	26.0	985	9	CG262551	CG262551 OG3AZA3TV
35	900	25.9	666	7	CF438196	CF438196 EST674541
36	899	25.9	697	5	BQ850117	BQ850117 QGB11L12
37	897	25.8	699	5	BQ849576	BQ849576 QGB10807
38	893.5	25.7	968	9	CG323044	CG323044 OG3AY44TV
39	882.5	25.4	778	8	CC005901	CC005901 PUGJY90TB
40	881	25.4	675	7	CK901981	CK901981 EST1084_C
41	876	25.2	656	5	BU101062	BU101062 WHE3361_B
42	875	25.2	866	7	CF519225	CF519225 CAP0003_I
43	862.5	24.8	968	6	CA195691	CA195691 SCEZAD107
44	851	24.5	683	5	BQ970363	BQ970363 QHB41N16
45	848	24.4	605	6	CA1134297	CA1134297 SCJFRT106

#### ALIGNMENTS

RESULT 1

CNS0A7F1

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Comment

CNS0A7F1 2158 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLRFB3C02 of Flowers and buds of strain col-0 of Arabidopsis  
thaliana (thale cress).  
BX822461 GI:42465936  
HTC; GSLT cDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsia.  
1 (bases 1 to 2158)  
Castell, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,  
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schacher, V.,  
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
Unpublished  
2 (bases 1 to 2158)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for protein  
 Sequences). 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full\\_length](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length)

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

## FEATURES

Location/Qualifiers  
 1..2158  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTPB32C02"  
 /tissue\_type="Flowers and buds"  
 /plasmid="pCMVSPORT\_6"  
 complement(1..2158)  
 /gene="At3g51895"

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,3e-272 Length: 2158  
 Score: 2416.00 Matches: 469  
 Percent Similarity: 85.02% Conservative: 93  
 Best Local Similarity: 70.95% Mismatches: 93  
 Query Match: 69.53% Indels: 7  
 DB: 3 Gaps: 4

US-10-762-049-18 (1-680) x CNS0A7F1 (1-2158)

Qy 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43  
 Db 62 ATGGGCACGGAGGACTACACATTCCTCAAGGAGCGGAGGAGCTTCACCGCCGCCACCAC 121  
 Qy 44 ---ValGluValProProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62  
 Db 122 ACGGTGGAGGCTCCACAACTCAACCGCTTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 181  
 Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProLysLysLysPhe 82  
 Db 182 ACTCTGTTCCAGACGACCCCTTTAGACAAATTTAGNACCMAAATGATCAAGAAATTT 241  
 Qy 83 MetLeuGlyLeuGlnPhePhePheProLeuPheGluTrpAlaProLysTyrThrPheGln 102  
 Db 242 GTGTAGGCCTCAAATATCTCTCCCGATTTCGAATGGGCACCGCTACAATCTCAAG 301  
 Qy 103 PheLeuLysAlaAspLeuLeuAlaGlyLeuThrLeuAlaSerLeuAlaLeuProGlnGly 122  
 Db 302 TTCTTCAAATTCAGATCTCATCGCGGAATCACCATCGCTAGTCTCGCCATCCCTCAGGGC 361  
 Qy 123 IleSerTyrAlaLysLeuAlaAsnLeuProProLeuLeuGlyLeuTyrSerSerPheIle 142  
 Db 362 ATCAGTTACGCCAAATCTGTAACTTGCCTCCCAATCTTGGCCCTTTATTCGAGTTTGT 421  
 Qy 143 ProProLeuLeuTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162  
 Db 422 CCGCCATTGTGTATACCGGTGCTAGGAGTTCAAGGAGCTTAGCGGTGGGAACGGTTGCG 481  
 Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182  
 Db 482 GTTGGCGTCTCTGTTGACAGTGCATGCTGAGCAAAAGATTGATGCTGAGAAAGATCCT 541  
 Qy 183 LysLeuTyr-LeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 202  
 Db 542 AAGCTTTACCAATTCACCTTGCTTTACCGCCACTTTTTCGCCGGGTCTTCGAGCCCTC 601  
 Qy 202 a-LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleG 222  
 Db 602 ATCTTGAATTTTCAGGTAGGTTTCATAGTGACCTTCTATCGCATGCAACGATAGTAG 661  
 Qy 222 LysPheMetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeuGlyL 242

Db 662 GATTTCATGGAGGAGCAGCAACGGTGGTGAGTCTGCAACAGAGATTAAAGGATATTTTAGGAA 721  
 Qy 242 euGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnT 262  
 Db 722 TTCAACATTTTCACAGACTCAACAATGTTATCACAGTCATGATTCATCAATTTTCACAAAA 781  
 Qy 262 hrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeu 282  
 Db 782 CACACAAGTGGAGATGGGAAGTGGCGTTTCATGGATGTGGTTTCAATTCATTCATCACA 841  
 Qy 282 erThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaPro 302  
 Db 842 CAACAGATATTTTCAGCATCAAGCAACCAAAATTCATTTGGTGGCGGATGGCTCAAT 901  
 Qy 302 euThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyV 322  
 Db 902 TGACACAGTGTATTTGGCAGTTCATGTGTTTTCATCACACACAGAGAGACATGGTG 961  
 Qy 322 alGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValP 342  
 Db 962 TTCAAGTTCATAGGGCCAGAGACAGAGTGGCTTCATCTCAAGT---TCTGATCTCATCT 1018  
 Qy 342 heValSerProTyrMetThrThrAlaValLysThrGlyLeuValValGlyIleIleSerL 362  
 Db 1019 TTTCTTTCACA-TTCATGTCCACAGCTGTCCACACCTGGCCACATCATCTGGCATCATTTGCTC 1077  
 Qy 362 euAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspG 382  
 Db 1078 TCGCTGAGGAGTAGCAGTGGGAGAGTTTGGCATGTTTCAAGAACTCAACATAGACG 1137  
 Qy 382 LysAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCyr 402  
 Db 1138 CGAACAAAGAGATGATAGCGTTTGGAAATGATCAACATCGTTGGTCTCTTCATCATCTGTT 1197  
 Qy 402 YrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrA 422  
 Db 1198 ACCTCAACACCGGACCATTTTCAAGTCCGCGAGTGAACACTACACCGGGTTCGAAGACCG 1257  
 Qy 422 laAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProL 442  
 Db 1258 CAATGTCCAAACATAGTATGGCGATTGCGTTATGTTTCACACTCTCTCTCTCTCACACCGC 1317  
 Qy 442 euPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyL 462  
 Db 1318 TTTTTCACATACACACCATCGTCTCTCTGCGCATCATCATATCCGCAATGCTCGGAC 1377  
 Qy 462 euIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValC 482  
 Db 1378 TCATTCAGCTATCAAGCTGCCATCTCTGGAAGTTGACAAAGTTCGACTTCTCTCGTCT 1437  
 Qy 482 YsMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI 502  
 Db 1438 GCATGAGCCCTACGTTGGGTCGTATTCGGCAGCTGTAGAGATTGGACTCTGCTGAGCGG 1497  
 Qy 502 leValIleSerValLeuArgValLeuLeuPhePheIleAlaArgProArgThrPheValLeuG 522  
 Db 1498 TGGCGCATCTATAGCAGGTTGTTGTTGTTGTCGAGGCCAAAACTCGCGGTGAAGG 1557  
 Qy 522 LysAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisV 542  
 Db 1558 GAAACATCAACAAACAGCATGATCTATAGAACACACTGAGCAGTACCCTCATCAAGAACCG 1617  
 Qy 542 alProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrL 562  
 Db 1618 TTCCTGGTATCTCATCTGGAGATTGATGCTCCCATCTACTTTCGTAATGCAGTTACT 1677  
 Qy 562 euArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyG 582  
 Db 1678 TCGCGTGAAGAATCATAAAGTGCATTGATGAAGAGAGAGAGAGTAAACAATCAGGAG 1737  
 Qy 582 LuThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerG 602  
 Db 1738 AGAGCAGCTTACAATATATATTACTCGATATGTCAGTGTGTTGGTAAATATCGACACAGCG 1797



```
Db 1269 TGCTACGTAAACAACAGGTGGTCTCTCCCGCTCTGCTGTAACACCAACGCGGCTGCAAG 1328
Qy 421 ThrAlaAlaSerAsnIlelleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db 1329 ACTGCCATGTCACACGATGATCGGCGCTGACTGTGATGCTACGCTGTCTTCCTCATG 1388
Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIlelleValSerAlaMetLeu 460
Db 1389 CCATGTTCTGTACACACCAACGTTGCTCTCGAGCGATCATCATCGCGCGGTGATC 1448
Qy 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db 1449 GGCCTGATCATGTTTCCCGCGGTGTACACATCTGGAAGATGACAAGATGGATTTCTG 1508
Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db 1509 GTGTGGCTTTCGCGGTTCGCGCGGTATCTTTCATCTCAGTCCAAAGAGCCCTTGCATA 1568
Qy 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
Db 1569 GCGTTGGTATCTATATTAGGTGTGTGATCGAGTACACAGGCCGAGATGATGGTT 1628
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db 1629 CAAGGGAAACATCAAGGGACTGATATTACAGAGACTGTCATCACTACAAGGAGGCCAA 1688
Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db 1689 AGAGTTCTCGGTTCTTGATCTTGGCCATTGGAAGCCGATAAATTCGCCAACTCCAAC 1748
Qy 561 TyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThr 580
Db 1749 TACTGATCAAGAGGATTAAGAAGTGGATA-----GAGNAGAACTTTTGACAGGAT 1802
Qy 581 GlyGluThrSerLeuGlnTyrValIlelleAspMetSerAlaValGlyAsnIleAspThr 600
Db 1803 AAACATACCTGAACTCCATTTTCATAATCTTGATCTGTGAGTCTGCTGCAATTTGACACA 1862
Qy 601 SerGlyIleSerMetLeuGluValValLysLysIleThrGluArgArgGluLeuGlnLeu 620
Db 1863 AGTGGCATAGCTTCTCATGTAATAAGAAATCAATAGAGAAACGTTGCTGTGGAGCTT 1922
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLys---SerLysPheGln 639
Db 1923 GTGCTTGTCAATCCAACCTGAGAGATCATGAGAGAAATCAACAGTCACAAACGAGGCTGAA 1982
Qy 640 AsnHisLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 659
Db 1983 AACTATTTTAGGCCAGATTCTGTATCTGACCACTGGCGAAGCAATC-----GCTTCA 2036
Qy 660 PheAsnLeuArgAlaSerLysThrAsnProLysLysAsp 672
Db 2037 CTTTCTGCACTTGCCAAGATGACAAAACCCCTAAATGGAT 2075

RESULT 3
LOCUS CL959635
DEFINITION OsIFC003265 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL959635
VERSION CL959635.1 GI:52373944
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
REFERENCE 1 (bases 1 to 2001)
AUTHORS Ma,L., Wang,C., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M., Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L., Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and
```

```
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
    Location/Qualifiers
        1..2001
            /organism="Oryza sativa (indica cultivar-group)"
            /mol_type="genomic DNA"
            /db_xref="taxon:39946"
            /clone_lib="Oryza sativa Express Library"
            /note="Oryza sativa exon trapped genomic sequences"
ORIGIN
Alignment Scores:
Pred. No.: 1,73e-201 Length: 2001
Score: 1813.50 Matches: 341
Percent Similarity: 73.10% Conservative: 121
Best Local Similarity: 53.96% Mismatches: 167
Query Match: 52.19% Indels: 3
DB: 9 Gaps: 3
US-10-762-049-18 (1-680) x CL959635 (1-2001)
Qy 43 GlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db 106 GAGGTGAACCTGTCCGGGCGCGCCGTTCGCGGAGAGCTGTGTTCGACCTCGCGGAG 165
Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82
Db 166 ACGTTCTTCCCGAGACGCCGTTCGCGGGTTCGCGCGCTCCGCCGCGAGCGCG 225
Qy 83 MetLeuGlyLeuGlnPhePhePheProIlePheGluTyrAlaProLysTyrThrPheGln 102
Db 226 TGTGCGCGGTCAAGTACTTCTCCGCGCTGGACTGGTCCCGCTACCGCTCGAC 285
Qy 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 286 AAGTTCAAGTTCGACTCTCTCGCGCATCACCATCGCCAGCTCGCCATCCGCGAGGC 345
Qy 123 IleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIle 142
Db 346 ATCAGCTACGCCCGCTCGCCAACTCCCTCCAATCATCGGCTCTATTTCGAGCTTCGTG 405
Qy 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 406 CCGCGTTGATGATGCGCGGTTCGGAGCTCCAACTCGCGCGCGAGCAACCCG 465
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
Db 466 GCGGCGTCTGCTGCTGCGCTCATCATCGAGCGAGGTGCGCGCGAGCAACCCG 525
Qy 183 LysLeuTyrIleHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 526 CAGCTGTACCTGCAGCTCTTTCACCGCGCTTCTTCACCGCGCTCTTCAGACCCGG 585
Qy 203 LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleLeuGly 222
Db 586 CTAGGCGCTTTAGGTAGGCTAATAGTGGATTCTCTCGCGGTTCGAGCATCACCGG 645
Qy 223 PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeu 242
Db 646 TTCATGGGCGGCACGCGCATGATCATCATCTGACGAGTTCACAGGCGCTCTCTCGGGGATG 705
Qy 243 GluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThr 262
Db 706 AAGCACTTCACCACCAAGACAGACATCATCTCCGCTCTGCTGCACCTACCATTACCGG 765
```

```
QY 263 HisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSer 282
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
766 CACGAGTGGAAATGGCAGAGTCAGATTCTCGGCATATGCTTCTCTTGTCTTGTCTGTCG 825
QY 283 ThrArgTyrPheSerLysArgProArgPhePheTrpValSerAlaMetAlaProLeu 302
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
826 AGTAAGCCTCGAAGAAAGAACTGCCAAGTTGTTCTGGGTGCGGCCATTCGCCATTTC 885
QY 303 ThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyVal 322
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
886 ATGGTTGCTGTCAGGAGCATCTTCGCTTCTCGTCAAGGTCGACGAGCATGGAATT 945
QY 323 GluValIleGlyGluLeuLysGlyLeuAsnProSerLeuThrAsnLeuValPhe 342
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
946 CCGATAGTCGGTACCTGGAATAAGGATCAATCCCTGTCCTATTCGGCAGCTAACGTTTC 1005
QY 343 ValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeu 362
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1006 CAGTCCAGGCACATGAGACGCGCAGTGAAGCTGGTCTTGTCTGGGATCCTAGCACTG 1065
QY 363 AlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGly 382
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1066 GCAGAGGAATAGCGTCGCGCGAAGCTTGGCGATGCTCAAGAACGACGACATCGACGGC 1125
QY 383 AsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr 402
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1126 AACAAGAGAGATGATCGCGTTTGGTATCATGAACATCCCGGTTCTTTTCACCTCTGCTAC 1185
QY 403 LeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAla 422
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1186 CTACGACAGCGCGGTCTCAAGTCGCGGTGAACCTTCCACGCGGGTGCAGAGCGCG 1245
QY 423 AlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuProLeu 442
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1246 ATGTCGAACGTGTGATGCGGTGTCATCATGCTGCTGTCTGCTGTCGACCCCTC 1305
QY 443 PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu 462
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1306 TTCAAAGTACACCCGCTGTGTGGGCTCTCTCCATCATCTGCTCGCATGATCGGCTC 1365
QY 463 IleAspTyrGluAlaAlaIleHleLeuPheLysValAspLysPheAspPheValValCys 482
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1366 GTCAAGTCAAGAGGTTCTGCCACCTCTACAAGTCGACAGTTTCGACTTCGATCTGC 1425
QY 483 MetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle 502
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1426 ATGTCGCTCTCTCGCGCTGCTCTTCTTCCACCATGTCGACGCTCAGCGCATCCGTA 1485
QY 503 ValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGly 522
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1486 GGCTTGTGCTGTCAGGGGCTGTCTGACGTGCGCAGGCTGCCACCTGCAAGCTCGGG 1545
QY 523 AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVal 542
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1546 AACATAGCAGGAGCAGACCTTCCGCGAGTCGACCAATACCCCAAGCAGGAGCATC 1605
QY 543 ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu 562
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1606 CTGGCATCTCTGCTGTGACGCTCGGCTCTCCCATCTACTTCTGTCATGCGGGTTACCTG 1665
QY 563 ArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGlu 582
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1666 CGAGAAAGGATTTTGATGGGTGAAGATGAGATAACCTGTGCAAGAGCGTCGGGCAC 1725
QY 583 ThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly 602
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1726 ---GATTGCAATACCTGGTCTTGATGTTGGTGGTGTCTACTTCTGTGCAACTCTCGGA 1782
QY 603 IleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeu 622
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1783 GTTGGGATGCTACTAGAAAGTACACAGAGGCTCGAACGAGAGGGATCAGCATAGTCTCTG 1842
```

```
QY 623 ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1843 ACGAATCCGAGGCTAGAGGTGACAGAGAAGCTGCTGCTGTGATACGTCAGGACATC 1902
QY 642 LeuGlyLysTrpIleTyrLeuThrValGluAlaValGlyAlaCysAsnPheAsn 661
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1903 TTAGGGATGAGTGGGTCTCTCTGACGGTCAAGGACGCCATCAGCGGTGCTCATACGGC 1962
QY 662 LeuArgAlaSerLysThrAsnProLysLysAspGlu 673
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1963 CTCAGATATCCAGA---AATAAGGAGAGACGAA 1995

RESULT 4
CNS09YVW 2221 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTPGH702F01 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX832405
VERSION BX832405.1 GI:42455111
KEYWORDS HTC; GSLT_cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE 1 (bases 1 to 2221)
AUTHORS Castellì,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpetti,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2221)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castellì
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EP/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
source
Location/Qualifiers
1..2221
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTPGH702F01"
/tissue_type="Hormone Treated Callus"
/plasmids="pCMVSPORT 6"
complement(1..2221)
/gene="At5g19600"

gene
ORIGIN
Alignment Scores: 6.12e-194 Length: 2221
Pred. No.: 1750.00 Matches: 334
Score: 72.13% Conservative: 119
Percent Similarity: 53.18% Mismatches: 158
Best Local Similarity: 50.36% Indels: 18
Query Match: 3 Gaps: 4
DB:
```



```
US-10-762-049-18 (1-680) x CNS09YVW (1-2221)
QY 44 ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGluThr 63
DB 127 GTAAATTTCTCGACCCGAGAGCTTTGGTTCAAAATTTAAGTCAAAATGTAAGAAACC 186
QY 64 PhePheProAspAspProLeuArgGlnPhePheLysAsnLysPro-----AlaSer 79
DB 187 TTCTTCCCGATCATCAATCAAAACCGATATCGCAAGAGCCAAACCGTTTGGTCAAAAACA 246
QY 80 LysLysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluThrPalProLysTyr 99
DB 247 AAGAAGTTA-----TTGGAGTATTTTGTGCCAATCTTTGAGTGGCTCCCAAGTAT 297
QY 100 ThrPheGlnPheLysAlaAspLeuAlaGlyIleThrIleAlaSerLeuAlaIle 119
DB 298 GATATGCAAAACTAAGTATGATGCTCGCCGGAATCACCATCAGCTTGTGCT 357
QY 120 ProGlnGlyLeuSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSer 139
DB 358 CCTCAGGCATCAGTATGCTAAGCTCGCCAGTATCCCTCCCATCATAGCCCTATATTC 417
QY 140 SerPheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGly 159
DB 418 AGTTTGTGCGCGCTTTTGTGATGCGTGTTCGGAAGTTTCGAACAATCTGCGAGTGGGA 477
QY 160 ThrValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsn 179
DB 478 ACHGTGCGAGCATGCTTTTGTGATCGTGTGAGACATTCGGTGAGAGATGATAAAAAAT 537
QY 180 GluAspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPhe 199
DB 538 GAG--CCTGAACCTCTACCTTCATTTGATCTTCCACGCCCACTCTCATCACCGGTCTATT 594
QY 200 GlnAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThr 219
DB 595 CAATTTGTCTATGGGATTTCTTGGGATTTGGGATATTTGGTGATTTCTGTGCACACTCGAC 654
QY 220 IleIleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIle 239
DB 655 ATACCGGCTTCATGGCGGTACCGCATCATCTTCTCTACACAGCTTAAGGCAATC 714
QY 240 LeuGlyLeuGlnHisPheThrHisGlyAlaAspIleSerValMetArgSerValPhe 259
DB 715 TTTGGACTAGTCCATTTTACGCACAAACCGATGTGATCAGTTCTCCACTCAATCTTA 774
QY 260 ThrGlnThrHisGluThrArgThrGluSerAlaValLeuGlyCysValPheIlePhePhe 279
DB 775 GACAAACCGGCTGAGTGGAGTGGCAAGTACCTTGGCTGGTCTGTCTTCTGTATTT 834
QY 280 LeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePheThrValSerAlaMe 299
DB 835 CTACAATCCACTCGATACATAAGCAAAAGATACCCAAAGCTATTTTGGGGGTATCAGCAAT 894
QY 299 tAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLys 319
DB 895 GGGTCCAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 954
QY 319 sHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAs 339
DB 955 ACACGGCATTCCTACTTGGGCTTTAAAGAAAGGACTAAATCCCTCTTCTTATTCACCT 1014
QY 339 nLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyL 359
DB 1015 -TTGAATTTGATCCCAAGTATCTAGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1073
QY 359 eIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAs 379
DB 1074 CATCGCTTTGGCCGAAGGATACCAATTTGGAAGAACCTTTTGTGTAATGAAGAACGAACA 1133
QY 379 nIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheTh 399
```

```
DB 1134 AACGACCGGAAACAAAGAGATGATGACATTTGGTCTTATGAACGTAATGGCTCTCTCAC 1193
QY 399 rSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCys 419
DB 1194 TTCCTTGCTATTTGACACACGGGCCATTTTCAAGACACAGAGTGAACACTACACGCGGGAC 1253
QY 419 sLysThrAlaAlaSerAsnIleMetSerLeuAlaValMetLeuThrLeuLeuPheLe 439
DB 1254 AAAGACACCGATGTGCAATGTAATGGCGCTTTGGCATGATGCTTGTCTCTCTCTCT 1313
QY 439 uThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMe 459
DB 1314 ACGCGCTCTATTACAGCTACACACCATAGTTGGTCTCTCAGCCCATCATTTATGTCAGCCA 1373
QY 459 tLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPh 479
DB 1374 GTTGGGTCTCATCACTACGAGAGATGTACCATCTCTTCAAGGTTGACAAAGTTTGATTT 1433
QY 479 eValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVa 499
DB 1434 CCTCGTCTGATGTCCGCTTTTTCGGTGTTCATTTCTTAGCATGGACTACGGTCTCAT 1493
QY 499 lIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPh 519
DB 1494 CATCTCCGTTGGATTTCTATCGTAAGAGCTTTGTGTACGTTGCAAGGCCCTCCACTTG 1553
QY 519 eValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAl 539
DB 1554 TAAATTTGGAGAGATACCAACTCGGTTATGTTTAGACATAGACAGTACCTTGCCTC 1613
QY 539 aLysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAl 559
DB 1614 CGAGGAGATGCTGTGTATATCATTTCTTCAATAGGCTCTCCAGTCTTTTTCGCAATAG 1673
QY 559 aSerTyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluGluArgIleLysAl 579
DB 1674 CACTTACGTCGCGTAACCGATCTTGAAGTGGATCTCGTGAACCTGAA----- 1722
QY 579 aThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAs 599
DB 1723 -----GCCATCGAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1769
QY 599 pThrSerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlu 619
DB 1770 CATGACCGGATGGAACACTACTCGAAATTCAGAGAATCTTGGATCAAAAACATCAA 1829
QY 619 nLeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGl 639
DB 1830 GATGTGTATTAATAATCCAAAGATTTGAAAGATGATGTTGTCTCCATTTCTGT 1889
QY 639 nAsnHisLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAs 659
DB 1890 GGAGAGATAGGAAAGAGTATATGTTTATTCGATGACGATGACGATGCAAGCATGCCG 1949
QY 659 nPheAsnLeuArgAlaSerLys 666
DB 1950 ATTTAATCTTACCACCACCAAG 1971

RESULT 5
CL969864 1812 bp DNA linear GSS 21-SEP-2004
LOCUS OsIFCC019054 Oryza sativa Express Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL969864.1 GI:52394347
VERSION GSS.
KEYWORDS Oryza sativa (indica cultivar-group)
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pharbitaceae; Oryzaeae; Oryza.
REFERENCE 1 (bases 1 to 1812)
AUTHORS Ma, L., Wang, C., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
```

Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,  
Wong, G.K.S., Deng, X.W. and Wang, J.  
An analysis of transcriptional regulation of the rice genome and  
its comparison to Arabidopsis  
Unpublished (2004)

**JOURNAL  
COMMENT**

Department of Bioinformatic  
Beijing Institute of Genomics  
Chinese Academy of Sciences, Beijing 101300, China  
Tel: 86-10-80481559  
Tel: 86-10-80488676  
Fax: 86-10-80488676  
Email: chenchen@genomics.org.cn  
Rice genomic sequence.  
Class: exon-trapped.

## FEATURES

```

1. .1812
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/notes="Oryza sativa exon trapped genomic sequences"

```

## ORIGIN

Alignment Scores:	
Pred. No.:	1812
Score:	1.31e-190
Percent Similarity:	1720.50
Best Local Similarity:	71.20%
Best Local Similarity:	54.21%
Query Match:	49.51%
DB:	9
	Indels: 29
	Gaps: 3
US-10-762-049-18 (1-680) x	CL969864 (1-1812)

[illegible]

Qy	241	GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr	260
Db	532	GGCATCGCTCCACTTACCACGAGATGGGCGCTCGCTCCGCTCATGGCTCCGTCATCCAC	591
Qy	261	GlnThrHisGlnLysArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeu	280
Db	592	CACACCAAGAGGTGGTGGTGGCAGACGATCCTGATGGCGGTGGTCTTCCCTGGGTGCTCGT	651
Qy	281	LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTyrValSerAlaMetAla	300
Db	652	CTCAGCGGAGGATGTGAGTATGAAATGCCAAAGCTTTCTGGGTTCAGCTTGCTGTCT	711
Qy	301	ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis	320
Db	712	CCCTTGGCATGTGTATCGTCTCTACCCCTGCTTGTTCCTGTTCAAGCTCAGAAACAT	771
Qy	321	GlyValGluValIleGlyGluLeuLysGlyLeuAsnProSerLeuThrAsnLeu	340
Db	772	GGCATTAGCATAAATGGCGACGTCAAATGTGGCTCGAATGCCCTTCATGGGATAAACTA	831
Qy	341	ValPheValSerProTyrMetThrThrAlaValIysThrGlyIleValValGlyIleIle	360
Db	832	CTCTTTGATCCTCAATATTTGGCGCTCAGTGTAAAGACTGGCCCTGTCCCGGAATCATC	891
Qy	361	SerLeuAlaGluGlyIleAlaValAlGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle	380
Db	892	TCCTTAACGGAAGAGTGGCGAGTTGGTAGACATTTGCTTCACTCAAGGACTACCAAGTA	951
Qy	381	AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer	400
Db	952	GATGGTAAACAAAGAGATGATGGCCATAGGATTCATGAACATTTGGGCTCATGTATCATCA	1011
Qy	401	CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys	420
Db	1012	TGCTATGTACTACAGGAGCATTTCTCTGTCTGTGTGACCAACACATGCTGTGTGCAAG	1071
Qy	421	ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr	440
Db	1072	ACTGCCATGTCAAAATGTGATCATGGCAGTCACTGTCTATGGTCACTGTTGTTTCCTCATG	1131
Qy	441	ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu	460
Db	1132	CCACTGTTGTGTACACACCGAATGTTGTTCTTGGAGCAATCAATATGCTGCTGTGATT	1191
Qy	461	GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal	480
Db	1192	GGGCTCATTGATCTCCCTGCTGTATACACATCTGGAAGATGGACAAAGATGGATTTCCTT	1251
Qy	481	ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle	500
Db	1252	GTGTGCTATGTGCAATTTGCTGTGTAAATCTTTCATCTCAGTCCAACAAAGCCCTTGCAATA	1311
Qy	501	AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal	520
Db	1312	CGCGTTGGTATATCAATATTCAGGGTACTGTGTGCAAAATCACAAGGCCAAAGATGTGATT	1371
Qy	521	LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys	540
Db	1372	CAGGGNACATCAAGGCACGTGATATTATCCGAAATCTTCATCAGTACAAAGGACGCTCAA	1431
Qy	541	HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer	560
Db	1432	AGAGTTCCTGGGTTCTTAATCCTGACATTTGAGGCTCCTATAAATTTGCAACACCAAC	1491
Qy	561	TyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThr	580
Db	1492	TACCTAAATGAAAGGATTTAAAGATGCATA-----GAAGAAGAAAGATTCTGCAGGAACC	1545
Qy	581	GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr	600
Db	1546	AAACAAAGGAACTCCATTTTCGTAATTTTGGATCTGTGAGCTCTCCCTGCAATTTGATACA	1605
Qy	601	SerGlyLeuSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeu	620

```

Db      1606 AGTGGCATATGTTCTCTCGACTTGAAGAAATCAACAGAGCATGCACTAGAGCTT 1665
QY      621 ValLeuValAsnProValSerGluValMetLysLeuAsnLysSerLys---PheGln 639
Db      1666 ATACTTGTGTAATCAACTGAGAGGTCATGGAGAGATACAACAGCGAATGATGCTCAT 1725
QY      640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAla 657
Db      1726 GGTCAITTCAAATCAGATAGTCGTACTTAACCACTGGGGAGCAGTGGCTTCA 1779

RESULT 6
CL961102
LOCUS   CL961102
DEFINITION
Oryza sativa (indica cultivar-group)
ORIGIN
Oryza sativa (indica cultivar-group)
VERSION
Oryza sativa (indica cultivar-group)
KEYWORDS
Oryza sativa (indica cultivar-group)
SOURCE
Oryza sativa (indica cultivar-group)
ORGANISM
Oryza sativa (indica cultivar-group)
REFERENCE
1 (bases 1 to 1959)
AUTHORS
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.
TITLE
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL
Unpublished (2004)
COMMENT
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source
location/Qualifiers
1..1959
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN
Alignment Scores:
Pred. No.: 6.63e-188 Length: 1959
Score: 1698.00 Matches: 328
Percent Similarity: 72.61% Conservative: 120
Best Local Similarity: 53.16% Mismatches: 164
Query Match: 48.86% Indels: 6
DB: 9 Gaps: 2

US-10-762-049-18 (1-680) x CL961102 (1-1959)
QY      42 HisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLys 61
Db      82 TACAATGTCGCGCGCGCGGAGAGAACCTCTCTCGCGGAGTTCGCCGCGCAGGTGAAG 141
QY      62 GluThrPhePheProAspAspProLeuAsgGlnPheLysAsnLysProAlaSerLysLys 81
Db      142 GAGACGCTTTCTCCGACGAGCGAGTCGCGGTACAGACGACGCGGAGTCCAGGAAG 201
QY      82 PheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPhe 101
Db      202 CTGTGGCTCGCCCTGAGCAGCGCTTCCCGGTTCGAATGGGGCAGACAATACACCTC 261
QY      102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121

```

```

Db      262 GCCAAGTTCAAGGGCGCACTCTATTCCGCGCTCAACCTTTGCCAGCGCTCGTCATACCTCAG 321
QY      122 GlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPhe 141
Db      322 GACATCGGCTACGCGAGCTTGTAACTGCCACACAGAGATTGGGCTGCACAGTAGCTTC 381
QY      142 IleProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal 161
Db      382 GTCCCGCCATCATATAGCTCTGTATGGGCACCTCAAGGGAGCTAGCAATGGTCCAGTG 441
QY      162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db      442 GCGCTCATCTACTGCTGTCTCTCTCTCCAGGAGGAGATTGACTCAAGAAGAAC 501
QY      182 ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db      502 CCGTAGATTACAGCGGCTCGCTTCACAGGACCTCTTTTGCAGGAGTTCACAGGGCG 561
QY      202 AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221
Db      562 GCGCTTGGTTTCTGCAGGCTAGGGTTTCATCATAGCGTTCCTGTCTCATGCTGCCATCATC 621
QY      222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db      622 GGATTCATGGCGCGCGCCCATCCATTGCTCTTCAGCAGCTTAAAGGCTTCTCTTGA 681
QY      242 LeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln 261
Db      682 ATTGCAAACTTCACCAAGAGACTGACATCATCTCGTTCATCAATCATCTGCGGGAAT 741
QY      262 ThrHis-----GluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePh 279
Db      742 GTTCACCATGGGGCAATGGAACTGGCAGACAATAATTGATCGGAGCATCATTTTGGCAAT 801
QY      279 eLeuLeuSerThrArgTyrPheSerLysArgProArgPhePheThrValSerAlaMe 299
Db      802 CCTCTGGT- TGCCAA-----GGCAAGAAGAACAAAGAGCTCTTCTGGGTCCCCAGCAAT 854
QY      299 tAlaProLeuThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLys 319
Db      855 TGACACCATCTATTCGGTGATCATTTCACTTTGTTGTTCTACATCATCTGCTGTGCACAA 914
QY      319 sHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAs 339
Db      915 ACAAGCGGTGCAATTTGTAAGAAACGTCAGAAAGGCATCAATCCACCTTCAGTAGCCT 974
QY      339 nLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyL 359
Db      975 GATATTTTCTACTGGCCCATACTTGTCTCAAGGATTTCAAAATTTGGAGTAGTAGTGAAT 1034
QY      359 eIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAs 379
Db      1035 GATAGACCTTACGGAAGCGATTGCGATTGGAGAACATTTGCTGGATTGAACGATTACCA 1094
QY      379 nIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheTh 399
Db      1095 GATAGATGGAAACAAAGAAATGTTGCTCTAGAACCATGAATGTGGTTCGTTCAATGAC 1154
QY      399 rSerCysTyrLeuThrThrGlyProPheSerSerAlaValAsnTyrAsnAlaGlyCy 419
Db      1155 GTCTTGTCTATATAGCCACAGGTGGTTTGGCAGATCAGAGTCAATTCGATCGCTGGAGG 1214
QY      419 sLysThrAlaAlaSerAsnIleMetSerLeuAlaValMetLeuThrLeuLeuPheLe 439
Db      1215 TAAACACCAATGTCCAATATTTGTTATGTCACTGTAGTATTGCTTGGCTCTCTGGAT 1274
QY      439 uThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMe 459
Db      1275 CACTCCATTTGTTCAAGTACACGCCCAATGCCACCATTTCTTCCATCATCATATCAGCAGT 1334
QY      459 tLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPh 479
Db      1335 GCTTGGCCTATTGTAATCAGCCTACCTTATCTGGAAGGTTTGATAGTTGGACTT 1394

```

```
Qy 479 eValValCysMetSerHlaTyrlleGlyValValPheGlySerValGluIleGlyLeuVa 499
Db 1395 CATGGCGTGTGGGGCATCTCTGGAGTAATATTTCATCTGTGGAGTATGGCTTGCT 1454
Qy 499 lIleAlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPh 519
Db 1455 CATTGGGTTGTATATCATCTAATCAAGTTCTGTCTCATGTAAACGGCCAGGACAGC 1514
Qy 519 eValLeuGlyAsnIleProAsnSerValIleTyArgAsnValGluHisTyArgGlnAsnAl 539
Db 1515 TTTACTTTGGCAACCTTCCAAAGAAGATTATCTATAGAAATGTTGAACAATATCCAGAAGC 1574
Qy 539 aLysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyPheAlaAsnAl 559
Db 1575 TACCAAGGTCCAGGGATGTAATGTGAAGAGTGGACTAGCAATATACTTCAAAACTC 1634
Qy 559 aSerTyLeuArgGluGluIleThrArgTTrpIleAspGluGluGluGluArgIleLysAl 579
Db 1635 CAATTATGTTAAAGAAAGATGCTGAGATGCTGAGAGATGAGAGAGAACATCAAAAGGA 1694
Qy 579 aThrGlyGluThrSerLeuGlnTyValIleIleAspMetSerAlaValGlyAsnIleAs 599
Db 1695 ACAGAAGTTACCAAAATTCAGTTCTGATTGTTGACCTATCTCTGTTAAATGATTTGA 1754
Qy 599 pThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgGluLeuGl 619
Db 1755 CACAAGTGAATCCATCCCTTCAAGAGTGTGTTGAGGACACTTGAAAAGGCCAGATTCA 1814
Qy 619 nLeuValLeuValAsnProValSerGluValMetLysLysLysLysLysLysLysPheGl 639
Db 1815 GCTGATTTTCCCAATCTCGGGCGGCTGTGATCCAAAGACTCCGGTCAGCAAAATTCAC 1874
Qy 639 nAsnHisLeuGlyLysLysTrpIleTyLeuThrValGluGluAlaVal 655
Db 1875 AGAGCTCATGTTGAGAAAGATATGCTTGACAGTTGTTGACGCCGTA 1923

RESULT 7
CNSOAILZ
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT
```

genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full length  
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.

FEATURES  
source  
1..2321  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTPGH82E06"  
/tissue\_type="Hormone Treated Callus"  
/plasmid="pCWSFORT\_6"  
1..2321  
/genes="At5g10180"

ORIGIN  
Alignment Scores:  
Pred. No.: 2,04e-161 Length: 2321  
Score: 1474.00 Matches: 300  
Percent Similarity: 68.47% Conservative: 130  
Best Local Similarity: 47.77% Mismatches: 185  
Query Match: 42.42% Indels: 14  
DB: 3 Gaps: 8  
US-10-762-049-18 (1-680) x CNSOAILZ (1-2321)

Qy 48 ProProGlnProPhePheLysSerLysTySerLysGluThrPhePheProAsp 67  
Db 246 CACCTAGCCCTGGGATGAG---CTCAAAAGACAGTCAAGGCTCTTTCCTA----- 236  
Qy 68 AspProLeuArgGlnPheLysAsn-----LysProAlaSerLysLysPheMetLeu 84  
Db 297 ACCAAAGCCAAAAGTTCAAGTCACTTCAAAACAGCCCTTCCAAAACAAATCTCTCT 356  
Qy 85 GlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyThrPheGlnPheLeu 104  
Db 357 GTCCTCCAAAGCCATTTTCCCAATCTTCGGTGGCAGAAACTATAAATCACCATGTTTC 416  
Qy 105 LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer 124  
Db 417 AGAAGCATCTCATGGCTGGTTTAAACCTCGTAGCTCTGCAATTCGCGAGCAGCATTTGGT 476  
Qy 125 TyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTySerSerPheIleProPro 144  
Db 477 TATGCAACTCTTGCAAGGCTTGATCTCAATATGGCTATATACGAGTGTGGTACCACCA 536  
Qy 145 LeuIleTyAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGly 164  
Db 537 TTGATATATGCATTGATGGGAGCATCAAGAGAGATAGCAATCGGACCGGTGGCTAGTA 596  
Qy 165 SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeu 184  
Db 597 TCTTCTTATATCTTCAATGTTTGAGAAACTCATCGATCCAGAAACAGATCCCTTGGGA 656  
Qy 185 TyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuGly 204  
Db 657 TACAAGAAACTGTCTCTAACCAACACTTCTTCGCGGGATCTTCCAAGCTCTCTTCGGT 716  
Qy 205 LeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMet 224  
Db 717 TTATTTCAGTTTAGGGTTCTTGGTGGATTTCTGTCGACGAGCCATATGGGTTCATG 776  
Qy 225 GlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis 244  
Db 777 GGTGGTGCAGCCATGTTAATTGGACTCCCAACAGCTTAAAGGTTGTTGGTATCATTAC 836  
Qy 245 PheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln---ThrHis 263  
Db 837 TTCACCACCAACACATGACATGCTCTCTCTCTTCGAGCTGTCTGGAGTCTTCTGCAACA 896  
Qy 264 GluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThr 283











```

Db      620 AACATTGGGATCACTCACCTTGTCTACCTCACCACCGGTCCGTCTCGAGGTGGCC 679
Qy      413 ValAsnTyrAsnIaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaVal 432
Db      680 GTCAACTACAAACGCGGCTGCAAGACGCGGATGCGAAGCGTATCATGTGCGTGGGGTG 739
Qy      433 MetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSer 452
Db      740 ATGATCAAGCTGCTGTTCTTGACCGCGCTGTTCACATACACGCGCGTGGTGGTGTGTCG 799
Qy      453 AlaIleIleValSer 457
Db      800 GCGATCATCATGTGCG 814

RESULT 11
CVI30479
LOCUS   X9P02a03 Populus stem seasonal library Populus deltoides cDNA, mRNA
DEFINITION
ACCESSION CVI30479
VERSION    CVI30479.1 GI:51874399
KEYWORDS  EST.
SOURCE    Populus deltoides
ORGANISM  Populus deltoides
REFERENCE 1 (bases 1 to 828)
AUTHORS   Park, S. and Han, K.-H.
TITLE     Gene expression profile during seasonal growth cycle in poplar tree
JOURNAL   Unpublished (2003)
COMMENT   Contact: Kyung-Hwan Han
          Department of Forestry
          Michigan State University
          126 Natural Resources, East Lansing, MI 48824-1222, USA
          Tel: 517 353 4751
          Fax: 517 432 1143
          Email: hanky@msu.edu.

FEATURES
source   Location/Qualifiers
1..828
/organism="Populus deltoides"
/mol_type="mRNA"
/strain="ILL-129"
/db_xref="taxon:3696"
/tissue_type="stem"
/dev_stage="1 year old"
/clone_lib="Populus stem seasonal library"

ORIGIN
Alignment Scores:
Pred. No.: 1.08e-120 Length: 828
Score: 1122.00 Matches: 217
Percent Similarity: 91.18% Conservative: 31
Best Local Similarity: 79.78% Mismatches: 22
Query Match: 32.29% Indels: 2
DB: 7 Gaps: 1

US-10-762-049-18 (1-680) x CVI30479 (1-828)
Qy      353 ThrGlyIleValValGly-----IleIleSerLeuAlaGluGlyIleAlaValGlyArg 370
Db      11 ACCGACATATGCCCGGAATTTCGGCCATTACGGCGGGGAAGGATAGCAGTAGGAAGA 70
Qy      371 SerPheAlaMetTyrIlysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGly 390
Db      71 AGTTTGGCATGTTCAAGATTACCATATAGATAGGCAACAAGAGATGATTGCAATTGGG 130
Qy      391 ThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArg 410
Db      131 ACCATGAACATTTGTAGGCTCTTGCACCTCTTGTCTATCTCAACAGGCCCATTCGCGCA 190
Qy      411 SerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeu 430

```

```

Db      191 TCGGCTGTGNAATTCAATGCAGGATGCAAGACAGCAGTATCGAATATTGTCATGGCGTTG 250
Qy      431 AlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValVal 450
Db      251 GCAGTCATGGTCACGTTGTTATTTCCTTAAGCCCATTTGTCATTACACTCCACTTGTGGTT 310
Qy      451 LeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHis 470
Db      311 CTATCTCTCTATTATTATCTCTGCCATGCTAGGTCTTTATAGATTATGAAGCAGCAATCCAT 370
Qy      471 LeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValVal 490
Db      371 CTTTGGACTGTCGACAAAGTTGACTTCATCGTGTGCAAGCGCATATGCTGGTGTGGTT 430
Qy      491 PheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeu 510
Db      431 TTTTGCAGTGTGCGAGATTGGCTTAGTCATCGCGCTTGGCAATTTCTTTGCTCAGATTACTT 490
Qy      511 LeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyr 530
Db      491 CTCTTCGTGCAAGACCAAAACATTATTCTTTGGAAACATTCCTCAATTCGATGATCTAC 550
Qy      531 ArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIle 550
Db      551 AGAAATGTTGAACAGTATACGAACAACAGCAGCGTTCCCGGTGTTCTCATCTCGAGATT 610
Qy      551 AspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrrIle 570
Db      611 GATGCTCCTACTCTTTCGAAATGCAAGTACTTAAAGAAAGAGGATTGCGAGGTGGTT 670
Qy      571 AspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrValIleIle 590
Db      671 GATGAAGAGAGAGAACAGTTAAATCTTCAGGAGAAACAGCCTCGAGTATGTTACTA 730
Qy      591 AspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLys 610
Db      731 GACATCGGAGCGGTAGTAACTTACACACAGTGGGATTGGCATGCTTCGAGGAAGTAAG 790
Qy      611 LysIleThrGluArgArgGluLeuGlnLeuValLeu 622
Db      791 AAAGTAATGGACAGAGGAAGTAAGTCAAGCTTGTCTTG 826

RESULT 12
CVI07436
LOCUS   GR_Eb0038K08.r GR_Eb Gossypium raimondii cDNA clone GR_Eb0038K08
DEFINITION
ACCESSION CVI07436
VERSION    CVI07436.1 GI:48806122
KEYWORDS  EST.
SOURCE    Gossypium raimondii
ORGANISM  Gossypium raimondii
REFERENCE 1 (bases 1 to 831)
AUTHORS   Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
          Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
          Wing,R.A.
TITLE     Global assembly of Cotton ESTs
JOURNAL   Unpublished (2004)
COMMENT   Contact: Rod A. Wing
          Arizona Genomics Institute
          The University of Arizona
          Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
          Tel: 520 626 9595
          Fax: 520 621 1259
          Email: http://genome.arizona.edu
          Plate: 0038 row: K column: 08.
          Location/Qualifiers
1..831
/organism="Gossypium raimondii"
/mol_type="mRNA"

GR_Eb0038K08.r GR_Eb Gossypium raimondii cDNA clone GR_Eb0038K08
3', mRNA sequence.

```

```
/db xref="taxon:29730"
/clone="GR_Eb0038K08"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/notes="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."
```

ORIGIN

Alignment Scores:  
Pred. No.: 1.64e-115 Length: 831  
Score: 1078.00 Matches: 201  
Percent Similarity: 88.45% Conservative: 44  
Best Local Similarity: 72.56% Mismatches: 31  
Query Match: 31.02% Indels: 1  
DB: 7 Gaps: 0

US-10-762-049-18 (1-680) x CO107436 (1-831)

QY 312 valTyrPheThrHisAlaGluLysHisGly-ValGluValIleGlyGluLeuLysLysG1 331  
Db 2 GTTATCTGACGATCGAGAAAACATGGCCCTTCAAGTGATTGGGCACCTTGAAGAAAG 61

QY 331 YLeuAenProProSerLeuThrAenLeuValPheValSerProTyrMetThrThraLaVa 351  
Db 62 GTTGAATCCACCTCTGCATCAGCCTGGTGTGGTGGCCCACTTGATGGCGGCTAT 121

QY 351 llySthrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSe 371  
Db 122 AAAACCGGGATCGTCATTGGCATCATTTGSCCTTGTGAAGGAATTCAGTCGGAAGAAC 181

QY 371 rPheAlaMetTyrLysAenTyrAenIleAspGlyAenLysGluMetIleAlaIleGlyTh 391  
Db 182 CTTTGGCATGTTCAAGAACTATCACATTGATGGGAACAAAGAAATGATCGCCTTCGGGAT 241

QY 391 rMetAenValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSe 411  
Db 242 GATGAACATGTGGGCTCATGCACCTCTCTGTACTTAACTGCAGGACCAATTTCAAGAAC 301

QY 411 rAlaValAenTyrAenAlaGlyCysLysThrAlaAlaSerAenIleIleMetSerLeuAl 431  
Db 302 TGCAGTGAACATTTAATGCTGGAAGCAGAAAGTGCAGTCTCCAACTTGTAAATGCAACAGC 361

QY 431 aValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValValle 451  
Db 362 AGTGATGTTTACGTTGTATTTCTAACACCATTTGTTCCATTACACTCTCTTGTGGTACT 421

QY 451 uSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLe 471  
Db 422 TTCCTCTAATATATAGCAGCCATGTTGGCTGCTGTGACTATGAAGCTGCAATCCACCT 481

QY 471 uPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValValPh 491  
Db 482 TTGGAATAATTGACAAATTCGACTTCTCATCTGCTTGGGGGCATATCTCGGTGTGTCCT 541

QY 491 eGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuLe 511  
Db 542 CGGAACGCTGTGAGATTGGTCTAATCATTTTCGATCATCTGTTTCATTGCTAAGGATTATTC 601

QY 511 uPheIleAlaArgProArgThrPheValLeuGlyAenIleProAenSerValIleTyrAr 531  
Db 602 CTTTGGCGGAGCCAGAACGATGTTTTAGGCATATATCCGAACCTCTGGGATCTATAG 661

QY 531 gAenValGluHisTyrGlnAenAlaLysHisValProGlyMetLeuIleLeuGluIleAs 551  
Db 662 AAGTAGGACCAATACCAATCGCAACAGTGTTCAGGGATTCCTCATACTGCAATAGA 721

QY 551 pAlaProIleTyrPheAlaAenAlaSerTyrLeuArgGluArgIleThrArgTrpIleAs 571  
Db 722 CGCCCCCTGTTTCTTTGCCAATGCAAGCTATCTTAAGAGAAGGATCTCGAGATGGATTTA 781

QY 571 pGluGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyr 587  
Db 782 TGAAGAGGAAACACAGGTTAAATCTGCAGAGAGAACGACCTTACACTAT 830

RESULT 13  
LOCUS CO102483  
DEFINITION GR\_Eb0029P05.r GR\_Eb Gossypium raimondii cDNA clone GR\_Eb0029P05  
3', mRNA sequence.

ACCESSION CO102483  
VERSION CO102483.1 GI:48801169  
KEYWORDS EST.  
SOURCE Gossypium raimondii  
ORGANISM Gossypium raimondii  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 706)  
Kim H., Yu Y., Kudrna D., Hatfield J., Stum D., Mueller C.,  
Udall J.A., Rapp R.A., Wendel J.F., Rao K., Soderlund C. and  
Wing R.A.  
Global assembly of Cotton ESTs  
Unpublished (2004)  
Contact: Rod A. Wing  
Arizona Genomics Institute  
The University of Arizona  
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: http://genome.arizona.edu  
Plate: 0029 row: P column: 05.

FEATURES  
Location/Qualifiers  
source  
1..706  
/organism="Gossypium raimondii"  
/mol\_type="mRNA"  
/db\_xref="taxon:29730"  
/clone="GR\_Eb0029P05"  
/tissue\_type="floral"  
/dev\_stage="3 to +3 DPA"  
/lab\_host="DH10B"  
/clone\_lib="GR\_Eb"  
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:  
EcoRV; Library made by Invitrogen with RNA supplied by  
Wendle lab. Directional cloned into NotI-EV. Colonies  
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:  
Pred. No.: 2.52e-108 Length: 706  
Score: 1016.00 Matches: 192  
Percent Similarity: 92.31% Conservative: 24  
Best Local Similarity: 82.03% Mismatches: 18  
Query Match: 29.24% Indels: 0  
DB: 7 Gaps: 0

US-10-762-049-18 (1-680) x CO102483 (1-706)

QY 303 ThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyVal 322  
Db 3 ACAACGGNNAITCTTGGAAAGTCTACTCGTTTACTTGACCCCATGCTGAAAAACACCGTGT 62

QY 323 GluValIleGlyGluLeuLysLysGlyLeuAenProProSerLeuThrAenLeuValPhe 342  
Db 63 GATGTGATTGGAACCTTGAAGAAAGGTTGAATCCACCTCTTTTGGGGAITTCGTATTT 122

QY 343 ValSerProTyrMetThrThraLaValLysThrGlyIleValValGlyIleSerLeu 362  
Db 123 ACCTCTCCCTATTTTACAACAGCTGTGAAGACTGCGCATGATCATCTGCTCTTT 182

QY 363 AlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAenTyrAenIleAspGly 382  
Db 183 GCTGAAGGAATAGCAGTGGGAGAGCTTTGCAATGTTTCAAGAACTACAACATTGATGGG 242

Qy	383	AsnLysGluMetIleAlaIleGlyThrMetAenValValGlySerPheThrSerCysTyr	400
Db	243	AACAAAGAAATGGTATGGGACCATGAACATTGTTGGTTCCTGCTTTCTTGGCTAT	302
Qy	403	LeuThrThrGlyProPheSerArgSerAlaValAenTyrAenAlaGlyCysLysThrAla	422
Db	303	CTCACACGGGGCCATTCTTCGATCAGCTGTTAACTTCATTCGGGATGTAACAACACA	362
Qy	423	AlaSerAenIleIleMetSerSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeu	442
Db	363	ATGTCGAACAGTGTAAATGGCCATTCAGCTTATGTTTCACATTGCTGTTCTTAACACCAT	422
Qy	443	PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu	462
Db	423	TTCATTATACATCCCTCGTAGTGTGTCTGCAATTAATAATCTCTGCAATGCTTGGCCCTC	482
Qy	463	IleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCys	482
Db	483	ATTGATTACGAGGCTGCATTCATCTTTGGAAAGTCGATAAATTCGATTGCTGTGTGHT	542
Qy	483	MetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle	502
Db	543	ATGGGTGCATATATCGGTGTTGTTTTCGCGAGTGTGGAGGTCGATTAGTCATAGCGGTT	602
Qy	503	ValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGly	522
Db	603	GCTATCTCTGTCTCGAGATCTGCTGTGTGTGCGAGACCGAAGACCTTCATTCTTGGGA	662
Qy	523	AsnIleProAenSerValIleTyrArgAenValGluHisTyr	536
Db	663	AACCTTCGAATTCAACCATCTATAGGAGTGTGGAGCAATAT	704
RESULT 14	CK256297	783 bp mRNA linear EST 30-JUL-2004	
LOCUS	EST73934	potato callus cDNA library, normalized and full-length	
DEFINITION	Solanum tuberosum cDNA clone POC529 5' end, mRNA sequence.		
ACCESSION	CK256297		
VERSION	CK256297.1	GI:39813277	
KEYWORDS	EST.		
SOURCE	Solanum tuberosum (potato)		
ORGANISM	Solanum tuberosum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
TITLE	asterids; lamids; Solanales; Solanaceae; Solanum.		
JOURNAL	1 (bases 1 to 783)		
COMMENT	Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B. Generation of ESTs from potato callus tissue Unpublished (2003) Other ESTs: EST739935 Contact: Robin Buell The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA Email: potato-acrv@tigr.org Clones can be requested from the University of Arizona Genomics Institute via http://genome.arizona.edu/orders/ Seq primer: Art TAG GTG ACA CTA TAG. Location/Qualifiers 1. 783 /organism="Solanum tuberosum" /mol_type="mRNA" /cultivar="Kennebec" /db_xref="taxon:4113" /clone="POCCS29" /tissue_type="callus" /lab_host="DH10B-Tona" /clone_lib="potato callus cDNA library, normalized and full-length" /note="vector: pCMWSPORT6.1; Site_1: EcoRI; Site_2: NotI; supplier: RNA was isolated from Solanum tuberosum var. Kennebec callus tissue grown on solid media."		
FEATURES	source		
ORIGIN			

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae; Allium.

## REFERENCE

1 (bases 1 to 785)

## AUTHORS

Havey, M.J., Cheung, F., Van Aken, S., Uterback, T. and Town, C.D.  
Expressed Sequence Tags from a normalized library of mixed onion

## TITLE

tissues (Allium cepa)

## JOURNAL

Unpublished (2003)

## COMMENT

Contact: Havey MJ

Department of Horticulture

USDA-ARS and University of Wisconsin

1575 Linden Drive, Madison, WI 53706, USA

Tel: 608-262-1830

Fax: 608-262-4743

Email: mjhavey@facstaff.wisc.edu

TIGR sequence name ACAEC64TR. For more information:

http://haveylab.hort.wisc.edu

Seq primer: CAG GAA ACA GCT ATG ACC.

## FEATURES

source

1..785

Location/Qualifiers

/organism="Allium cepa"

/mol\_type="mRNA"

/cultivar="Red Creole (bulbs), unknown (callus), Ebano &

Texas Legend (roots)"

/db\_xref="taxon:4679"

/clone="ACASC64"

/tissue type="Callus, roots, and young bulbs"

/clone\_lib="normalized cDNA library of onion"

/note="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site\_1:

EcoRV (5'); Site\_2: NotI (3'); Equal molar amounts of mRNA

from callus, roots, and young bulbs were combined to

synthesize the library. Normalization to enrich for

low-copy transcripts was performed by proprietary

techniques of Invitrogen."

## ORIGIN

## Alignment Scores:

Pred. No.:	2.28e-106	Length:	785
Score:	1000.00	Matches:	193
Percent Similarity:	82.29%	Conservative:	30
Best Local Similarity:	71.22%	Mismatches:	24
Query Match:	28.78%	Indels:	24
DB:	7	Gaps:	3

US-10-762-049-18 (1-680) x CF438329 (1-785)

QY	228	AlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHis	247
Db	2	GCAACTGTAGTTGTCTACAGCAGTGTGAAAAGTATTTGGGACTTCAGCATTTTACTACA	61
QY	248	GlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGlu	264
Db	62	TCCACTGATCTCTCTCTGTATGGAGTCTGTGTTGCAGTCTCACCAGGCATGCTGG	121
QY	265	-----TtpArgTtpGluSerAlaValLeuGlyCys	274
Db	122	CGTTGGAATATTCTAAGTAGTCCAGAAATGGAGATGGAGAGTGTGTTCTTGGATGT	181
QY	275	ValPheIlePheLeuLeuSerThrArgTyrPheSerIlyLysArgProArgPhePhe	294
Db	182	GGATTCTTATCTTCTCTTGAGCACTCGATTTTTCAGTAAAGAGGCCAAATTTATTT	241
QY	295	TtpValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPhe	314
Db	242	TGGGTGCTGCGACAGCTCCACTAACTCGGTCACTCTTGGCAGTCTACTGCTTACTTTC	301
QY	315	ThrHisAlaGlyLysHisGlyValGluValIleGlyLeuLysLysGlyLeuAsnPro	334
Db	302	ACCAAGCAGAAAACCATGGTGTCAAGTTATTTGTTATCTAAAGAAAGGGTAAATCCA	361
QY	335	ProSerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGly	354
Db	362	CCCTCTGTCAAAAAGCTGGTCTTCTCAGCACCTTATATTACGGTGTCTCTCAAAGCTGA	421

Search completed: August 31, 2005, 16:01:07  
Job time : 5829 secs

QY	355	IleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMet	374
Db	422	ATCATCACTGGGTATCATTTGCACCTTCTGCTGAAGCGCATAGCAGTAGGAAGGAGCTTTTCAATG	481
QY	375	TyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnVal	394
Db	482	TTCAAGAATTACCATCGATCGAAACAAAGAGATGATCGCGTTTGGGATGATGAACATG	541
QY	395	ValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsn	414
Db	542	GCCGGTCCATCAGCATCATGCTACTTAAACACCGGACCACTTCGAGGACAGCTGTAAAC	601
QY	415	TyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeu	434
Db	602	TACAATGCAGGATGCAAGACGCAATGTCAAATGTTGGTATGCGCATAGCAGTGATGATC	661
QY	435	ThrLeuLeuPheLeuThrProLeuPheHisTyrThrPro---LeuValValLeuSerAla	453
Db	662	ACGCTACTATTTTAAACACCGCTTTTCCACTACACTCCCTCTTTGTGGTCCCTTGCCTCT	721
QY	454	IleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLys	473
Db	722	ATTATTATATCGCAATGCTTGTCTTATCGACTACGAGCTGCTGCCCTC-----	775
QY	474	ValAspLysPheAspPheValValCysMetSer	484
Db	776	-----TGCAATGTCG	784

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2005, 11:07:04 ; Search time 425 Seconds  
(without alignments)  
9428.804 Million cell updates/sec

Title: US-10-762-049-17  
Perfect score: 2449  
Sequence: 1 gcacgagctagctgcacat.....aaaaaaaaaaaaaaaaaaaaa 2449

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues  
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2449	100.0	2449	US-09-720-317A-17	Sequence 17, Appl
2	765.2	31.2	1981	US-09-720-317A-3	Sequence 3, Appl
3	720.6	29.4	2067	US-09-720-317A-15	Sequence 15, Appl
4	655.4	26.8	2279	US-09-720-317A-1	Sequence 1, Appl
5	640.4	26.1	2311	US-09-720-317A-19	Sequence 19, Appl
6	324	13.2	780	US-09-720-317A-7	Sequence 7, Appl
7	216.2	8.8	1240	US-09-720-317A-5	Sequence 5, Appl
8	205.6	8.4	2022	US-09-720-317A-21	Sequence 21, Appl
9	118	4.8	493	US-09-720-317A-13	Sequence 13, Appl
10	117.4	4.8	484	US-09-720-317A-9	Sequence 9, Appl
11	90.4	3.7	2441	US-09-785-381-4	Sequence 4, Appl
12	82.6	3.4	4927	US-09-949-016-5627	Sequence 5627, Ap
13	79.8	3.3	4113	US-09-785-381-2	Sequence 2, Appl
14	77.6	3.2	2487	US-09-248-796A-6643	Sequence 6643, Ap
15	76.4	3.1	7218	US-08-232-463-14	Sequence 14, Appl
16	64.4	2.6	2692	US-09-614-221A-222	Sequence 222, App
17	63.8	2.6	2913	US-09-795-927-6	Sequence 6, Appl
18	63.8	2.6	3749	US-09-795-927-8	Sequence 8, Appl
19	58.2	2.4	510	US-09-720-317A-11	Sequence 11, Appl
20	58	2.4	758	US-09-270-767-470	Sequence 470, App
21	58	2.4	758	US-09-270-767-15752	Sequence 15752, A
22	57.8	2.4	3061	US-09-570-842-1	Sequence 1, Appl
23	57.8	2.4	3061	US-08-700-576-1	Sequence 1, Appl
24	56	2.3	1773	US-09-902-540-7746	Sequence 7746, Ap
25	56	2.3	4854	US-09-902-540-768	Sequence 768, App
26	53.4	2.2	2832	US-09-949-016-4516	Sequence 4516, Ap
27	53.4	2.2	8774	US-09-949-016-16258	Sequence 16258, A

28	53.2	2.2	320	4	US-09-270-767-27504	Sequence 27504, A
29	53.2	2.2	904	4	US-09-270-767-11855	Sequence 11855, A
c 30	52.8	2.2	1116	4	US-09-252-991A-7460	Sequence 7460, Ap
31	52.8	2.2	1791	4	US-09-252-991A-7241	Sequence 7241, Ap
32	51	2.1	761	4	US-09-270-767-1826	Sequence 1826, Ap
33	51	2.1	761	4	US-09-270-767-17108	Sequence 17108, A
34	51	2.1	1767	4	US-09-252-991A-7184	Sequence 7184, Ap
35	47.4	1.9	1971	4	US-09-875-811-9	Sequence 9, Appl
36	47.4	1.9	1992	4	US-09-875-811-5	Sequence 5, Appl
37	47.4	1.9	2040	4	US-09-875-811-1	Sequence 1, Appl
38	47.4	1.9	2420	4	US-09-875-811-13	Sequence 13, Appl
39	47.4	1.9	7218	1	US-08-232-463-14	Sequence 14, Appl
c 40	45.4	1.9	1141	4	US-09-806-708B-22	Sequence 22, Appl
41	45.4	1.9	1143	4	US-09-540-236-913	Sequence 913, App
c 42	45.4	1.9	31147	4	US-09-596-002-25	Sequence 25, Appl
43	45.2	1.8	1466	4	US-09-620-312D-914	Sequence 914, App
44	45.2	1.8	2858	4	US-09-949-016-5349	Sequence 5349, Ap
45	45.2	1.8	2882	1	US-08-424-567-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-09-720-317A-17  
; Sequence 17, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 17  
; LENGTH: 2449  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-720-317A-17

Query Match	100.0%;	Score 2449;	DB 4;	Length 2449;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 2449;	Conservative 0;			
QY	1	GCACGAGCTAGCTCGCACATTAAGTTATATACACATATTTGCTTGTCTTAGAAATACTAT	60	
Db	1	GCACGAGCTAGCTCGCACATTAAGTTATATACACATATTTGCTTGTCTTAGAAATACTAT	60	
QY	61	TATTGAAGATATGGGAGGTAGATATAGTACCTTTGGGCATGAACACTTTGAGAG	120	
Db	61	TATTGAAGATATGGGAGGTAGATATAGTACCTTTGGGCATGAACACTTTGAGAG	120	
QY	121	AGTGACCAAGTCGAGGTTCCACCCGACGCGTTTTTCAAGTCTCTAAAGTACTCTTT	180	
Db	121	AGTGACCAAGTCGAGGTTCCACCCGACGCGTTTTTCAAGTCTCTAAAGTACTCTTT	180	
QY	181	GAAGGAGACTTTCTTCCCTGATGACCTTTGAGGCAGTTTCAAGAACAGCCAGCTTCCAA	240	
Db	181	GAAGGAGACTTTCTTCCCTGATGACCTTTGAGGCAGTTTCAAGAACAGCCAGCTTCCAA	240	
QY	241	GAAGTTCAAGTCTGGCCTTCAGTTCTTCTTCCCAATTTTGAATGGGCTCCCAATACAC	300	
Db	241	GAAGTTCAAGTCTGGCCTTCAGTTCTTCTTCCCAATTTTGAATGGGCTCCCAATACAC	300	
QY	301	CTTTGAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATCC	360	
Db	301	CTTTGAGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATCC	360	

QY 361 TCAGGCGATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATCTTGGACTATATTCGAG 420  
DB 361 TCAGGCGATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATCTTGGACTATATTCGAG 420  
QY 421 CTTTATACCAACCAATTCATTTATGCCGATGATGGGTAGCTCGAGGGAATTCGGCAGTGGGAC 480  
DB 421 CTTTATACCAACCAATTCATTTATGCCGATGATGGGTAGCTCGAGGGAATTCGGCAGTGGGAC 480  
QY 481 TGTGGCGGTGGATCGCTCTCTGATGGGTTGATGTTGAGTAATGCGGTTGATCCCAATGA 540  
DB 481 TGTGGCGGTGGATCGCTCTCTGATGGGTTGATGTTGAGTAATGCGGTTGATCCCAATGA 540  
QY 541 AGACCCAAAGCTTTACTCCACCTCGCTTTACAGACTACATTTATTTGCTGGTGTTCATCA 600  
DB 541 AGACCCAAAGCTTTACTCCACCTCGCTTTACAGACTACATTTATTTGCTGGTGTTCATCA 600  
QY 601 GCGTGCCTTGGGTCTGTTAGGTTGGGTTGATCGTGGGATTTCTGTCACATGCAACCAAT 660  
DB 601 GCGTGCCTTGGGTCTGTTAGGTTGGGTTGATCGTGGGATTTCTGTCACATGCAACCAAT 660  
QY 661 AATAGGTTTCATGGGAGGAGCAGCCACGTTGGTGTCTGCGAGCAACTAAATTCGATTC 720  
DB 661 AATAGGTTTCATGGGAGGAGCAGCCACGTTGGTGTCTGCGAGCAACTAAATTCGATTC 720  
QY 721 TGGCCTTGACATTTTCAACCATGGAGCTGATATCATATCAGTGTGCGCTCTGTTTTCAC 780  
DB 721 TGGCCTTGACATTTTCAACCATGGAGCTGATATCATATCAGTGTGCGCTCTGTTTTCAC 780  
QY 781 CCAAACTCATGAGTGGAGGTGGGAAAGTCTGTTGTTAGGATGTCCTTCATTTTCTTCT 840  
DB 781 CCAAACTCATGAGTGGAGGTGGGAAAGTCTGTTGTTAGGATGTCCTTCATTTTCTTCT 840  
QY 841 CTTTACACAAAGATCTTCAGCAAAAAACGACCAAGGTTTTTTTTGGGTGTGACCAATGCG 900  
DB 841 CTTTACACAAAGATCTTCAGCAAAAAACGACCAAGGTTTTTTTTGGGTGTGACCAATGCG 900  
QY 901 GCCATTGACGTCGGTTATATTTGGGAAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGCA 960  
DB 901 GCCATTGACGTCGGTTATATTTGGGAAGTCTCTTGGTTTATTTTCACTCAGCCGAGAGCA 960  
QY 961 CGGTGTTGAAGTGATAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAATCT 1020  
DB 961 CGGTGTTGAAGTGATAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAATCT 1020  
QY 1021 GGTATTTGTGTCGCTTACATGACTACAGCTGTGCAAACTGGCATGTGCTGGCATCAT 1080  
DB 1021 GGTATTTGTGTCGCTTACATGACTACAGCTGTGCAAACTGGCATGTGCTGGCATCAT 1080  
QY 1081 ATCACTTGGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTTACAAAT 1140  
DB 1081 ATCACTTGGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTTACAAAT 1140  
QY 1141 TGATGCAACAAAGAGATGATAGCTATTTGGGACCATGAAGCTAGTTGTTTCTTTCACCTC 1200  
DB 1141 TGATGCAACAAAGAGATGATAGCTATTTGGGACCATGAAGCTAGTTGTTTCTTTCACCTC 1200  
QY 1201 TTGCTACCTCACAACAGGACCAATTTTCGGCTTCGGCTGTGAACATAAAGCTGGATGCAA 1260  
DB 1201 TTGCTACCTCACAACAGGACCAATTTTCGGCTTCGGCTGTGAACATAAAGCTGGATGCAA 1260  
QY 1261 GACAGCAGCTTCCAAATTTATATGTCATTTGCGAGTAATGTTGACATTTCTTCTCGAC 1320  
DB 1261 GACAGCAGCTTCCAAATTTATATATATGTCATTTGCGAGTAATGTTGACATTTCTTCTCGAC 1320  
QY 1321 ACCCTTGTTCATTTACATCCCTCGTGGTGTGCTATCAGCTATTTATCGTATCTGCAATGCT 1380  
DB 1321 ACCCTTGTTCATTTACATCCCTCGTGGTGTGCTATCAGCTATTTATCGTATCTGCAATGCT 1380  
QY 1381 TGGACTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTGT 1440  
DB 1381 TGGACTCATAGATTATGAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTTGT 1440  
QY 1441 GGTGTGCATGAGTGCCATACATTTGGCGGTGCTTTGGCAGTGTGTGAAATTTGGCTTAGTCAT 1500

DB 1441 GGTGTGCATGAGTGCCATACATTTGGCGGTGCTTTTGGCAGTGTGTGAAATTTGGCTTAGTCAT 1500  
QY 1501 AGCTATTGTAATATCTGTAATCTTGGGTAATCTTCTATTTATGCAAGGCAAGACATTCGT 1560  
DB 1501 AGCTATTGTAATATCTGTAATCTTGGGTAATCTTCTATTTATGCAAGGCAAGACATTCGT 1560  
QY 1561 TTTGGGCAACATTTCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAA 1620  
DB 1561 TTTGGGCAACATTTCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAA 1620  
QY 1621 ACATGCTTCTGGAATGCTAAATTTAGAGATTTGATGACCAATTTTACTTTTGGCAATGCCAG 1680  
DB 1621 ACATGCTTCTGGAATGCTAAATTTAGAGATTTGATGACCAATTTTACTTTTGGCAATGCCAG 1680  
QY 1681 CTTATTAAGAGAAAGGATCACAAGGTGGAATGATGAAGAAGAAAGAAATTTAAAGCTTAC 1740  
DB 1681 CTTATTAAGAGAAAGGATCACAAGGTGGAATGATGAAGAAGAAAGAAATTTAAAGCTTAC 1740  
QY 1741 AGGGAGAGCTAGTTTGCAGTATGTTATTAATTTGATATGATGCTGTTGAAACATTTGATAC 1800  
DB 1741 AGGGAGAGCTAGTTTGCAGTATGTTATTAATTTGATATGATGCTGTTGAAACATTTGATAC 1800  
QY 1801 AAGTGGAAATAGTATGCTTGAAGAGGTGCAAGAAGATTACAGAGAGAGAGCTACAGCT 1860  
DB 1801 AAGTGGAAATAGTATGCTTGAAGAGGTGCAAGAAGATTACAGAGAGAGAGCTACAGCT 1860  
QY 1861 TGTTTTGGTCAATCCTGTAAGTGAAGTGAAGAACTGAAGAACTGAAGTTCCTCAAAA 1920  
DB 1861 TGTTTTGGTCAATCCTGTAAGTGAAGTGAAGAACTGAAGAACTGAAGTTCCTCAAAA 1920  
QY 1921 TCATTTAGGGAAGAAATCGATCTATCTGCTGTTGAAGAGGCGTTGAGGATGCAACTT 1980  
DB 1921 TCATTTAGGGAAGAAATCGATCTATCTGCTGTTGAAGAGGCGTTGAGGATGCAACTT 1980  
QY 1981 CAATCTGCTGCAAGCAAAACGAAACCAAGAGATGAAGAGAGAGGTTGGAAACAATGT 2040  
DB 1981 CAATCTGCTGCAAGCAAAACGAAACCAAGAGATGAAGAGAGAGGTTGGAAACAATGT 2040  
QY 2041 GTGACTGAGTCAATATGCCCCAAGAGATTTCTAAATAACTCAAAAAGCTTTATTCGTTTTCGT 2100  
DB 2041 GTGACTGAGTCAATATGCCCCAAGAGATTTCTAAATAACTCAAAAAGCTTTATTCGTTTTCGT 2100  
QY 2101 CTTAGTAAATGTTACCACTACCAATGTCGCAATGTCGAATCTCAGCACTCAGCGGAGAGT 2160  
DB 2101 CTTAGTAAATGTTACCACTACCAATGTCGCAATGTCGAATCTCAGCACTCAGCGGAGAGT 2160  
QY 2161 TTTAAAGGCATAGGAAATGAAAGATGCAAGGCTCTTCTAAATTTCTCAACTCTGCATCCT 2220  
DB 2161 TTTAAAGGCATAGGAAATGAAAGATGCAAGGCTCTTCTAAATTTCTCAACTCTGCATCCT 2220  
QY 2221 TAGTTAGAGAAATCTCCTATGATAGGCTGTTGAAATAATCTTTACGATCATGCTTG 2280  
DB 2221 TAGTTAGAGAAATCTCCTATGATAGGCTGTTGAAATAATCTTTACGATCATGCTTG 2280  
QY 2281 ATATATATTTCAAGAGAAATGCTAGCAACACTCTCAGACACACTCTTTTGAACACATG 2340  
DB 2281 ATATATATTTCAAGAGAAATGCTAGCAACACTCTCAGACACACTCTTTTGAACACATG 2340  
QY 2341 TAAAGAGGTAAAGAGTGTGTTGCTAGCACTCTCCTCAATTTCAATTTGAAAGTAATTCCT 2400  
DB 2341 TAAAGAGGTAAAGAGTGTGTTGCTAGCACTCTCCTCAATTTCAATTTGAAAGTAATTCCT 2400  
QY 2401 ATGAGATTTTAAAAATTCCTTTGGAAAAAATAAAAAAATAAAAAA 2449  
DB 2401 ATGAGATTTTAAAAATTCCTTTGGAAAAAATAAAAAAATAAAAAA 2449

RESULT 2  
US-09-720-317A-3  
; Sequence 3, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:

APPLICANT: Stephen M. Allen  
APPLICANT: Saverio C. Falco  
APPLICANT: Catherine J. Thorpe  
TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
FILE REFERENCE: BB-1167  
CURRENT APPLICATION NUMBER: US/09/720,317A  
CURRENT FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/092,833  
PRIOR FILING DATE: 14-07-1998  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 3  
LENGTH: 1981  
TYPE: DNA  
ORGANISM: Zea mays  
US-09-720-317A-3

Query Match 31.2%; Score 765.2; DB 4; Length 1981;  
Best Local Similarity 65.8%; Pred. No. 4.3e-202;  
Matches 1112; Conservative 0; Mismatches 578; Indels 0; Gaps 0;

QY 313 GRAAGCTGACCTCATAGCTGGGCATCACCATCGCTAGCTTGGCCATTCCTCAGGGCATCAG 372  
DB 7 GGAGTCCGACCTGATCGCGGCATCACCATCGCCAGCTCGCCATCCCGAGGGCATCAG 66  
QY 373 TTATGCCAAGCTCGCCAAACCTCCCTCCAAATTCCTTGGACTATATTCGAGCTTTTATACCAC 432  
DB 67 CTACGCCAAGCTCGCCAAACCTCGCGCCGCTCGGCTCGGACTCTACTCGAGCTTCGTGCGGCC 126  
QY 433 ATTGATTATGCGATGATGGGTAGCTCGAGGGATTTGGGACGTGGGGACTGTGCGGTGG 492  
DB 127 GCTGGGTGACGCTGATGGGGAGCTCCAAGGACCTGGCGGTGGGGACGCTGGCGGTGGC 186  
QY 493 ATCGCTTTCGATGGGTTCGATGTTGAGTATGATGCGCTTGATCCCAATGAAGACCCAAAGCT 552  
DB 187 GTGCTGCTCATCAGCTTCATGCTCGGCGAGGAGTGTGCGCCACGAGAACCCCGTGTCT 246  
QY 553 TTACCTCCACCTCGCTTTTACAGCTACATTTATTTGCTGTGTGTTTTCAGGCTGCCCTTGGG 612  
DB 247 CTACCTGCACTCGCTTTACCGCCACCTTCTTCGCGGGCTCTTTCAGGCTCGCTCGG 306  
QY 613 TCTGTTAGTGGGTGATGCTGTGATTTCTGTGATGATGATGATGATGATGATGATGATGATGAT 672  
DB 307 CCTCCTCAGGTTGGGCTTCATGCTGACCTGCTGTGCGACGCGCGACGATCGTGGGGTTCT 366  
QY 673 GGGAGGACGACCAAGTGTGCTGCGACCACTAAATCGATTCCTTGGCTTGAGCA 732  
DB 367 GCGCGCGCGGCGACGGTGGTGTGCTGCGACGAGCTGAAGGGCATGCTGGGGCTCGTCCA 426  
QY 733 TTTCCACCATGGAGCTGATATCATATCAGTGTGCTGCTGTTTTCACCCAAACTCATGA 792  
DB 427 CTTACCACTCCACCGAGCTGCTCGCTCATGGAATCCGCTTCAGCGACACACCA 486  
QY 793 GTGAGGTGGGAAGTGTGTTAGGATGTGCTTCAATTTTCTTCTCCTCTTAGCACAAG 852  
DB 487 GTGGGCTGGGAGAGCTCTGCTGCGCTCGGCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCT 546  
QY 853 ATACTTCAGCAAAAACGACCAAGTTTGTGGGTGTCAGCAATGCGCGCCATTGACGTC 912  
DB 547 CTTTCATCAGCAAGAGCGCTCCCAAGCTGTTCTGGATCTCGCGGCGCGCGCTTGAACGTC 606  
QY 913 CGTTATATTGGGAAGTCTCTTGGTGTATTTTCACTCAGCGCGAGAACGACGGTGTGAAAGT 972  
DB 607 CGTCGTCTCGGAGGCTTCTGTGTACCTCAGCGACGCTGAAGAACACCGGCATCGAAGT 666  
QY 973 GATAGAGAACTGAAGAGGTTTGAATCCACATCACTCACAATCTGTTGTTGTTGTTGTTGTTGTT 1032  
DB 667 GATCGGTTACCTGAAGAAAGGCTGAATCCACCGTGGTGAACAAGCTGCAATTTCTCACC 726  
QY 1033 GCCTTACATGACTACAGCTGTCAAACTGGCATTTGCTTGGCATCATATCATCTTCTCGGA 1092  
DB 727 GGCCTACATGATGCTCGCGCTCAAGACTGGGATCATCGGGCGTCAITTGCCCTCGCCGA 786

QY 1093 AGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATATAAATTAACAATATTGATGGCAACAA 1152  
DB 787 AGGAATCGCGTGGGAGGAGCTTCGCCATGTTTCAAGAACTACCAATACGCGAACAACAA 846  
QY 1153 AGAGATGATAGCTATTGGGACCATGAACGTAGTGGTCTTTTCACTCTTGTACTCTAC 1212  
DB 847 GGAGATGATCGCGATCGGACGATGAACGCTCTGGGCTCGCTCAGCTGTGTCTACCTGAC 906  
QY 1213 AACAGGACCATTTTTCGCGTTTCGCGTGTGAATATTAACGCTGGATGCAAGACGACGCTTC 1272  
DB 907 CACGGGGCCCTCTCGCGCTCGCGCTGAACCTAACAACGCGGGTGCAGGACGCGCATGTC 966  
QY 1273 CAACATTAATGTCACTTGCAGTAATGTGTGACATTTGTTATTTCTTGAACAACCTTGTTC 1332  
DB 967 GAAACGTGGTCACTCGCTGGCGGTGATGCTCAACGCTGCTGTCTCCTGACGCGCTGTCCA 1026  
QY 1333 TTACACTCCCTGGTGGTCTATCAGCTATTATCTATCTCAATGCTTGGACTCATAGA 1392  
DB 1027 CTACACGCGCGTGGTGGTCTGTGCGCGATCATCTCTCCGATGCTGGGCTTGGTCCA 1086  
QY 1393 TTATGAAGCAGCATCCATCTATTAAAGGTTGACAAATTTGACTTTGTTGTTGCTGATGAG 1452  
DB 1087 CTTGCGGGCGCGCTGCACTGTGCGGCTGCAAGGTCGACTTCTGCTCTGCGCCG 1146  
QY 1453 TGCATACATTTGGCGTGGTCTTTGGCAGTGTGGAATTTGGCTTAGTATAGTATTGTAAT 1512  
DB 1147 CGCGTACCTGGCGTGGTCTTTCGCGAGCGTCAAGGTCGGCTGGTCTGCGCGTCCGCGT 1206  
QY 1513 ATCTGTACTTCGGGTACTTCTATTTTTCAGAGGCAAGGACATTCGTTTTCGGGCAACAT 1572  
DB 1207 CTCCCTGTCTCCGCTGCTGCTGCTGCGCGCGCCAGGACACGGTGTCTCGGCAACAT 1266  
QY 1573 TCCAAATCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCMAAACATGTTCTCTGG 1632  
DB 1267 CCGCGGACCATGCTGTACCGGAGGATGGAACAGTACGCGCGCGGCGAGACGCTGCCCG 1326  
QY 1633 AATGTAATTTAGAGATTTGATGCAACCAATTTTCTTTGCAATGCCAGCTATTTAAGAGA 1692  
DB 1327 CGTGTCTGTGCTGCGCTGCGAGCGCGCTCTACTTTCGCAACGCGAGCTACTGCGAGA 1386  
QY 1693 AAGGATCACAAGTGGTGTGATGAAGAAGAAGAAATTAAGCTACAGGGGAGACTAG 1752  
DB 1387 GAGGATCTCGCGGTGGTTCGACGAGGAGGAGCGCACCAAGAGCCAGGGCGAGATGGG 1446  
QY 1753 TTTGCAATGTTTAAATTTGATATGAGTGTCTTTCGAAACATTTGATACAAGTGGAAATAG 1812  
DB 1447 CGTGGGTACGTTTCTCGACTGGTGGCTCCATCGGTACCATCGACAGCGGGAGCGAG 1506  
QY 1813 TATGCTTGAAGAGGTGAAGAAGATTTACAGAGAGAAGAGCTACAGCTTTGTTTGGTCAA 1872  
DB 1507 CATGCTGGACGAGCTCAACAGTCTTGGACAGGAGGGAATGCAGATCGTGTGGCGAA 1566  
QY 1873 TCCTGTAAGTGAAGTGAAGAAGAACTGAACAAATCGAAGTTCCAAATCATTTAGGAA 1932  
DB 1567 CCGCGGACGCGAGATCATGAAGAAGCTGACAGCTCCAAGGTGCTGGAGCAGATCGGCCA 1626  
QY 1933 GAAATGGATCTATCTGACTGTTTGAAGAGCGCGTTCGAGCATGCCAACTTCAATCTACGTC 1992  
DB 1627 CGAGTGGGTGTTCCCGACGCTGGGCGAGCGGCTGGCTGCTGCGACTAGCTGCTGCACTC 1686  
QY 1993 AAGCAAAACG 2002  
DB 1687 GCACAAGCG 1696

RESULT 3  
US-09-720-317A-15  
; Sequence 15, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe



; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167

; CURRENT APPLICATION NUMBER: US/09/720,317A

; CURRENT FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; PRIOR FILING DATE: 14-07-1998

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 15

; LENGTH: 2067

; TYPE: DNA

; ORGANISM: Glycine max

; US-09-720-317A-15

	Query Match	29.4%;	Score 720.6;	DB 4;	Length 2067;
	Best Local Similarity	62.2%;	Pred. No. 1.1e-189;		
	Matches 1134;	Conservative	0;	Mismatches 689;	Indels 0; Gaps 0;
QY	143	CGGCCACAGCGGTTTTCAGTCTCTAAAGTACTCTTTGAAGGAGACTTCTTCCCTGAT	202		
Db	8	CCACACCCAGCACACTCCACAACTTAGGCCACAGAGTCTCCGAAATCTTCTCCAGAT	67		
QY	203	GACCTTTGAGGAGTTCAAGACAGCCAGCTTCCAGAAAGTTCATGCTTGCCCTTCAG	262		
Db	68	GACCTCTCCACCGTTTCAAGAACCAACTCGGCTTTAAAGAGTTCTCTCCGACTTCAG	127		
QY	263	TTCTTCTTCCCAATTTTGAATGGGCTCCCAATACACCTTTTCAGTCTTTGAAAGCTGAC	322		
Db	128	TATCTCTTCCCAATTTTCAGCTGGGCCCCAATCACTCAATCTTACCCTTCTCCGCTCTGAC	187		
QY	323	CTCATAGCTGGCAATCCATTCGCTAGCTTGGGCAATTCCTCAGGCGCATCAGTTATGCCAAG	382		
Db	188	CTCATCTCGGCTCACCATTGCGAGCTCGCCATTTCTCAGGGAATCAGTTATGCCAAG	247		
QY	383	CTCGCAACCTCCCTCCAACTCTTGACTATATTCAGAGCTTTATACCACTTGTATTTAT	442		
Db	248	CTTGCCAACTGCGCACTATTTTGGATATATTCAGAGTTTGTTCCTCCCATTTGATATAC	307		
QY	443	GCGATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGAGCTGTGGCGGTTGGATCGCTTCG	502		
Db	308	TCGCTGCTTGGAGTTCTAGACATCTTGGTGTGGACCTGTTCCATTGGCTTTGGTC	367		
QY	503	ATGGGTTCCAGTTGAGTAATGCGGTTGATCCCAATGAAGACCCAAAGCTTTACCTCCAC	562		
Db	368	ATGGGATCAATGTTAAGTGATAAAATTTCTTACACTCAAGAACCTATTCTCTATCTGGGA	427		
QY	563	CTGGCTTTACAGCTACATTTATTTGCTGGTGTGTTTTCAGCTGCTGCTGTTTAGG	622		
Db	428	TTGGCTTTACCCGCCACTTTCTTGTGCTGGTGTATTCGAAGCTTCTCTGGGTATTTAAG	487		
QY	623	TTGGGTTGATCGTGGATTTCTGTCACTGCAACCATTAATAGGTTTCATGGGAGGACA	682		
Db	488	CTAGGCTTCGTAATTCATTTCTGTGCAAGGCAACGCTGTTGGATTCAACAGCGGTGCT	547		
QY	683	GCCACGGTGTGTCTGAGGAACTAAATTCGATTTCTTGGGCTTGGAGCATTTACCCAT	742		
Db	548	GCCATTAATTTGTCTCACTGCAGAGCTGAAAGGTTTACTTGAATAGTGCATTTACCCAG	607		
QY	743	GGAGCTGATATCATATCAGTATGCGCTCTGTTTTCACCCCAACTCATGAGTGGAGTGG	802		
Db	608	AGATGCAAAATATTCAGTAACGATCTCTGTTTTCAGCAAGACACGAGTGGTCATGG	667		
QY	803	GAAAGTGTGTGTTAGGATGTGTCTTCATTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT	862		
Db	668	CAAAACCATTTCTTGGATTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	727		
QY	863	AAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGCGGCCATTCAGTCGCTTATTTG	922		
Db	728	TTGAGGAAACCAAACTTAATCTGGGTTTCAGCAGCTGCCCATTTGACATCAGTTATTCG	787		
QY	923	GGAAGTCTCTGTTTATTTTCACTCAGCGCGAAGCAGCGGTGTTCAAGTGTATGAGAA	982		
Db	788	TCAACCATTTTAGTCTTTCTCTGAGAAATAGACTCATCAATTTTCAGTTATTTGGGCAC	847		

RESULT 4

US-09-720-317A-1

; Sequence 1, Application US/09720317A

QY	983	CTGAAGAAGGGTTTGAATCCACCATCACTCAAAATCTGGTATTTTGTGTCSCCTTACATG	1042
Db	848	TTACCAAGGGAGTTAATCCACCATCCACCAACATGTTATATCTCAATGGTCTTACTTG	907
QY	1043	ACTACAGCTGTCAAAACCTGGCATTTGCTTGGCATCATATCACTTGGCGAAGGATAGCA	1102
Db	908	GGTCTTGCTATCAAAACCTGGCATCATACAGGGATCTTATCTCCTGGAAGGATTCGA	967
QY	1103	GTAGGAAGAAGCTTTGCAATGTATAAAATTTACAATATTTGATGGCAACAAAGAGATGATA	1162
Db	968	GTAGGGAGAACATTTGCTTCACTTAAGAACTACACAGGTGGATGAAACAAAGAAATGATG	1027
QY	1163	GCTATTTGGGACCATGAACGTAGTTGGTCTTTTACCTCTTGGTACTCACAACAGGACCA	1222
Db	1028	GCCATTTGCTTAATGAACATAGCTGGCTGCTGTTCTTCACTGTTATGTTTACAAACGGGATCC	1087
QY	1223	TTTTTCGGCTTCGGCTGTGAATATAAAGCTGATGATGATGATGATGATGATGATGATGAT	1282
Db	1088	TTTTCTCGATCGGCTGTAACTATAATGCTGGACAGACAGACAGTTTCAATATATATC	1147
QY	1283	ATGTCACCTTGCAGTAATGTTGACATTTGTTATTTCTGACACCCCTTGTTCATTTACATCCC	1342
Db	1148	ATGGCTGCAGCTGTTCTAGTGACACTTCTGTTTCTCATGCTCTTTTCTACTATACACA	1207
QY	1343	CTGTGGTGTCTATCAGCTATTTATGCTATCTGCAATGCTTGGACTCATAGATTTAGAACCA	1402
Db	1208	AATGTTGTCTTAGCGGCCATTTATCATCACTGCTGTGATTTGCTTAATAGATTTATCAATCT	1267
QY	1403	GCCATCCATCTATTTAAGGTTGACAAATTTTCACCTTTGCTGTCATCAGTGCATACATT	1462
Db	1268	GCAATATAATTTGGAAGGTTGACAACTTGATTTCTTGGCTGTTTGGCTCTCTTTTTT	1327
QY	1463	GCGCTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCTAGTCTATTTGTAATATCTGTACTT	1522
Db	1328	GGGTTCTGTTCAITTTTCACTGCTTTAGTCTTGGTATAGCGGTTATCATATCAGTCTCTC	1387
QY	1523	CGGTPACTTCTATTTTATTTGCAAGGCAAGGACATTCGTTTGGGCAACATTTCCAAATCT	1582
Db	1388	AAGATCTGCTTTCATGTCACCTCGACCAACACTTTTGGTGTGGGGAATATACAGGAACA	1447
QY	1583	GTGATATACCGAAATGTTGACACTATCAAAATGCAAAACATGTTCTCGGAATGCTAATT	1642
Db	1448	CAATATTTCCACACATAAACCATACAAAGAGCTTTAAGAGTTCTTCTCATTTCTCAT	1507
QY	1643	CTAGAGATTTGATGCAACCAATTTTCTTCCCAATGTCAGCTATTTAAGAGAAGGATCACA	1702
Db	1508	TTGGCTGTTGAGTCTCCAATCTATTTTCTAACTCAACTCACTTATCTTCAAGAAAGGATCTG	1567
QY	1703	AGGTGGATTTGATGAAGAAGAAAGATTTAAAGCTACAGGGAGACTAGTTTGCAGTAT	1762
Db	1568	AGATGGTTCGAGAAGAGAGAGCATATAAAGCTATAAAGCTATAATGGAGCTCCATTTGAAGTGC	1627
QY	1763	GTATAATTTGATATCAGTGTCTTGGAAACATTTGATACAGTGGAAATTAAGTATGCTGAA	1822
Db	1628	ATAATTTTAGACATGACAGCTGTCAAGCCACAGACAGAGTGGGCTTGACACTTTATGT	1687
QY	1823	GAGGTGAAGAAGATTTACAGAGAAGAGAGCTACAGCTTGTGTTTGGTCAATCTCTTAAGT	1882
Db	1688	GAACTTAGAAGATCTCGGAGAAGAGATCACTTGAGTTTGTGCTGGCAATCTCTGTTGGA	1747
QY	1883	GAAGTGATGAAGAACTGAACTGAATCGAAGTTCAAAATCATTTTAGGGAAGAAATGGATC	1942
Db	1748	AATGTGATGGAAAAATTCATAAGTCAAACTATTTTGGATTTCTTTGGATTTAAAGGAGTGC	1807
QY	1943	TATCTGACTGTTGAAGAGCGCGT	1965
Db	1808	TATCTCAGCTGGGAGAGCTGT	1830

; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 1

; LENGTH: 2279  
; TYPE: DNA  
; ORGANISM: Zea mays  
; US-09-720-317A-1

Query Match 26.8%; Score 655.4; DB 4; Length 2279;  
Best Local Similarity 60.4%; Pred. No. 1.6e-171;  
Matches 1120; Conservative 0; Mismatches 726; Indels 9; Gaps 2;

QY 121 AGTGCAACAGTCGAGGTTCCACGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTT 180  
DB 188 AGTGCAACAGTGGCGCGCCACCGCGCGGAGCACGGGAGCAAGATGAAGGTGAGGTT 247  
QY 181 GAAGGAGACTTCTTCCCTGATGACCTTTTGAGGCGAGTTCAAGAACAGCCAGCTTCCAA 240  
DB 248 GAAGGAGACTTCTTCCCGACGACCCGTTCCGGGCGTTCAAGGGGAGCGCGCGGGAC 307  
QY 241 GAAGTTTCATCTTGGCTTCAGTTCCTTCTTCCCATTTTTCGAATGGGCTCCCAAAATACAC 300  
DB 308 GCAGTGGCTCATGGCGGTGAGTACCTTCTTCCCATCTCTGGACTGGGTGCGGAGTACTC 367  
QY 301 CTTTTCAGTCTTGAAGCTGACCTCATPAGCTGGCATCACCATCGCTAGCTTGGCCATTC 360  
DB 368 CTTTGTGCTCTTCAAGTCCGACCTCGTGGCGGSCCTCACCATTTGCCAGCCTGCCATTC 427  
QY 361 TCAGGGGATCAGTTATGCCAAGCTCGGCAACCTCCCTCCAAATCTTTGGACTATATTCGAG 420  
DB 428 TCAGGGGATTTAGCTACGCGAAGCTGGCAAGCTTGGCCCTCCCATATCGGGCTGTATTCGAG 487  
QY 421 CTTTATACCAACCATTTGATTTATCGATGATGGGTAGCTGGAGGATTTTGGCAGTGGGAC 480  
DB 488 CTTTGTGCGCCGATGTTGACCGGTGACCGGTGCTGGGGAGCTCCGTAACCTGGCGGTGGGCC 547  
QY 481 TGTGGCGGTTGGATCGCTTCTGATGGGTTTCGATGTTGAGTAATTCGCGTTGATCCCAATGA 540  
DB 548 GGTGTGATCTCGTCTGATCATGATGGGTCATGCTGGCGAGCGCTGAGCCCTCACTGC 607  
QY 541 AGACCCAAAGCTTTACTCCACCTGGGTTTCACAGCTACATATTTTGTCTGGTGTGTTTCA 600  
DB 608 GGAGCGGACGCTGTTCTGAGCTGGGCTTTCACCTCCACCCTGTTCCGCGGGGCTGGTGCA 667  
QY 601 GGCCTGCTGGGCTGTTTGGTGGGTTGATCGTGGATTTTCTGTCATGCAATGCAACCAT 660  
DB 668 GGCCTCCCTGGGATCTTCAAGCTCGGCTTCGATCGACTTCCCTGTCGAAGGCGAGCT 727  
QY 661 AATAGGGTTTCATGGGAGGAGCAGCCACGGTGGTGTCTGTCAGCAACTAAATTCGATTTCT 720  
DB 728 GGTGGGTTTCATGGCGCGCGCCCATCATCGTGGCGCTGCAGCAACTCAAGGGGCTGCT 787  
QY 721 TGGCCTTGAGCATTTCAACCATGGAGCTGATATCATATCAGTGAATGGCTCTGTTTTCAC 780  
DB 788 GGGCATCGTCCACTTCACACCCAGATGGGCATCGTCCCAAGTCAAGGCTCCGCTCTTCCA 847  
QY 781 CCAAACTCATGACTGAGGTGGAAAGTGTGTTAGGATGTTCTTCAATTTCTTCTCT 840  
DB 848 CCACACGAGGAGTGGTGGGAGACGATTCCTCATGAGGCGCTGCTCTCTGCTCTCTCT 907  
QY 841 CCTTAGCACAAGATACTTCAGCAAAAAACGACCAAGGTTTTTTTGGGTGTCAGCAATGGC 900

DB 908 GCTGTGCGGAGGCATGTGAGCATCAGATGGCCAAAAGCTTTTCTGGGTTTTCGCGCTGCGC 967  
QY 901 GCCATTGACGTCCGTTTATATTGGGAGTCTCTTGGTTTATTTTCACTCAGCCCGAGAGCA 960  
DB 968 GCCCTTGGCATCGGTCAACATCTGACGCTGCTTGTGTTTCTTCTTCAAAGCTCAGAACCA 1027  
QY 961 CGGTGTTGAAGTGATAGGAGAACTGAAGAAGGTTTGAATCCACCACTCACTCAAAATCT 1020  
DB 1028 TGGCATCAGCATCAATGGGCAGCTCAAGTGGCGCTGAATCGCCCTCGTGGGACAAGCT 1087  
QY 1021 GGTATTTGTGTCCTTACATGACTACAGCTGTCAAACTGGCATTTGTGTTGGCATCAT 1080  
DB 1088 CCTGTTTGAACACGGGCTATTTAGGCGCTCAACATGAAGACTGGCTTGTCAACGGAATCAT 1147  
QY 1081 ATCACTTGGGAAGGAATAGCAGTAGGAGAAGCTTTGCAATGTATAAAAAATTTACAATAT 1140  
DB 1148 CTCACCTGACGGAAGGAATAGCGTTGGTTAGAACATTTGCTCTCACTCAAGGACTACAGAT 1207  
QY 1141 TGATGGCAAACAAAGAGATGATAGCTATTGGGACCATGAACGTAAGTTGTTCTTTTCACTC 1200  
DB 1208 AGATGGAAACAAAGGAGATGATGCCCATAGGGTTGATGAATGTTTGGGTCTCTGCACATC 1267  
QY 1201 TTGCTACCTCAACACGAGACCAATTTTTCGGCTTCGGCTGTGAACATATAACGCTGGATGCAA 1260  
DB 1268 ATGCTACGTAAACACAGGTGCGTTCTCCGCTCTGCTGTAAACCAACACGCGGCTGCAA 1327  
QY 1261 GACAGCGCTTCCAAACATTAATATGCTCACTTCAGCTAATGTTGACATGTTATTCCTGAC 1320  
DB 1328 GACTGCGCATGTCAAACGTGATCATGGCGCTGACTGTGATGGTCAAGCTCTGTTCTCAT 1387  
QY 1321 ACCCTTGTTCATATACATCCCTCGTGTGCTATCAGCTATTATCGTATCTGCAATGCT 1380  
DB 1388 GCCACTGTTGCTGTACACCAACCGTTGTCTCGGAGGATCATCATCGCCGCGGTGAT 1447  
QY 1381 TGGACTCATPAGATTTAGAGCAGCCATCCATCTATTTAAGTTGACAAATTTGACTTTGT 1440  
DB 1448 CGGCTGATCGATTTTCCCGCGGTGTACCAATCTGGAAGATGGAACAAGATGATTTCT 1507  
QY 1441 GGTGTGATGAGTGCATACATTTGGCGGTGCTTTCGAGTGTGAATTTGGCTTAGTTCAT 1500  
DB 1508 GGTGTGCGTTTTCGGGTTTGGCGGCTCATCTTCATCTCAGTCCAAAGAGGCTTCGCT 1567  
QY 1501 AGCTATTGTAATATCTGTACTTTCGGGTACTTCTATTTATGCAAGGCCAAGGACATTCGT 1560  
DB 1568 AGCGGTGGTATATCTATATTTAGGGTGTGATGAGATCAACAGGCCGAGATGATGCT 1627  
QY 1561 TTTGGGCAACATTTCCAAATTTCTGATATACCGAAATTTGTGAGACTATCAAAATGCAAA 1620  
DB 1628 TCAAGGGAAACATCAAGGGGACTGATATTTACAGAGACCTGTCATCACTACAGAGGCCCA 1687  
QY 1621 ACATGTTCTGGAACTGCTAATTTCTAGAGATTCATGCCCAATTTACTTTGCCAATGCCAG 1680  
DB 1688 AAGAGTTTCTGGGTTTCTGATCTTGGCCATTTGAAGCCGATTAACCTTCGCCAATCCAA 1747  
QY 1681 CTATTTTAAGAGAAAGGATCACAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTAC 1740  
DB 1748 CTACCTGATGAAGAGATTAAGATGGAT-----AGAGGAAGATCTTTTGAACAGGA 1801  
QY 1741 AGGGAGACTAGTTTTCAGATATGTTATAATGATATGAGTGTGTTGGAACAATTGATAC 1800  
DB 1802 TAAACATCTGAACCTCCATTTCTAATCTTGGATCTGTGAGCTGTTCTGCAATTTGACAC 1861  
QY 1801 AAGTGAATAGTATGCTTTGAAGAGGTGAAGAGATTACAGAGAGAGAGACTACAGCT 1860  
DB 1862 AAGTGGCATAGGTTCTCTCATTTGACATTAAGAAATCAATAGAGAAACGTTGGTCTGAGCT 1921  
QY 1861 TGTGTTTGGTCAATCTGTAAGTGAAGTGAAGAAACTGGAACAAATCGA---AGTTCCA 1917  
DB 1922 TGTGCTTGTCAATCCAACTGGGAGAGTATGAGAGAAATATCAACGTGCAACAGGCTGA 1981  
QY 1918 AAATCAATTTAGGGAAGAAATGGATCTATCTGATCTGTGTAAGAGGCGGTTGGAGCA 1972  
DB 1982 AAATCAATTTAGGCCAGATTGCTGTATCTGACCACTGGCGAAGCAATCGCTTCA 2036

## RESULT 5

US-09-720-317A-19

; Sequence 19, Application US/09720317A

; Patent No. 6696292

; GENERAL INFORMATION:

; APPLICANT: Stephen M. Allen

; APPLICANT: Saverio C. Falco

; APPLICANT: Catherine J. Thorpe

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167

; CURRENT APPLICATION NUMBER: US/09/720,317A

; CURRENT FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; PRIOR FILING DATE: 14-07-1998

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 19

; LENGTH: 2311

; TYPE: DNA

; ORGANISM: Triticum aestivum

US-09-720-317A-19

Query Match 26.1%; Score 640.4; DB 4; Length 2311;

Best Local Similarity 59.4%; Pred. No. 2.3e-167;

Matches 1104; Conservative 0; Mismatches 751; Indels 3; Gaps 1;

QY	122	GTGCAACAAGTCGAGGTTCCACCGCCACACCGCGTTTTTCAAGTCTCTAAAGTACTTTG	181
DB	187	GTGTACAAAGTGGCTATCCCTCCGAAGAACTTGGCCACAGAGTTTACAGAAACATTG	246
QY	182	AAGGAGACTTTCTCCCTGATGACCTTTGAGCAGTTTCAAGACAGGACGCTTCCAG	241
DB	247	AGGGAGACTTTCTTCCACGACAAACCGCTGCGTCAATGAGGGCCAAATCCGGACCGAGG	306
QY	242	AGTTTCATCGTTGGCTTCAGTTCTTCTTCCCATTTTCGAATGGCTCCCAATACACC	301
DB	307	AGTTTCATGAGGGCTGAGTTCTTGTTCCTATATTTGGTGGGTAGGATTTACAGT	366
QY	302	TTTCAGTTCTTGAAGCTGACCTCATPAGCTGGCATCACCATCGCTAGCTTGGCCATTC	361
DB	367	CTCAACAAGTCAAGGCGATCTGATTGCCGATTGAACCATCGAAGTCTCTGTATTC	426
QY	362	CAGGGCATCAGTTATGCCAAGCTCGCCAACTCCCTCCAAATCTTGGACTATATTCGAGC	421
DB	427	CAGGACATGGCTATTTCGAAGCTTGTCTAATCTGGATCCGAGTATGGGCTTTACTCCAGC	486
QY	422	TTTATACCAACCATTTGATTTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACT	481
DB	487	TTCAATTCCTCCATTGATCTATGCTGCAATGGGTAGCTCAAGGGATATAGCGATTGGTCCA	546
QY	482	GTGGCGGTTGGATCGCTCTTGATGGGTTTCGATGTTTGAGTAAATGCCGTTGATCCCAATGAA	541
DB	547	GTTCGTGTGGTTCTCTTTTGATAGTTTCACTTCTACAGCTGAGGTTGACCATGTCAA	606
QY	542	GACCCAAAGCTTTCCTCCAACCTGGCTTTCAAGCTACATATTTGCTGGTGTTCAG	601
DB	607	ACAAGGAGGAATACATGCGCTCGCTTTTCAAGCAACCTTCTTCCGCTGATCACTCAA	666
QY	602	GCTGCTTGGGTCTGTTTAGTTGGGTTGATCGTGGATTTCTGTCAATGCAATGCAATATA	661
DB	667	GCAGCCTTAGGATTTCTAAGGTTAGGATTCCTTATAGAGTTCTTGTCCGATGCTGGATT	726
QY	662	ATAGGGTTATGGGAGGAGCAGCAGCGGTGGTGTCTGCAGCAATATAAATCGATTCTT	721
DB	727	GTCCGATTATAGGGGAGCTGCCATTAATTTGCCCTGCAGCAGCTGAAATACGTGTTG	786
QY	722	GGCTTTGAGCATTTACCCATGGAGCTGATATCATATCAGTATGCGCTCTGTTCAC	781
DB	787	GGCATCGAACTTTTACAGGAAACCGACATAGTTTCTGTATGGAATCTGTCTGGAGA	846
QY	782	CAAACTCAT---GAGTGGAGGTTGGGAAAGTGCTGTGTAGGATGTCTTCATTTTCTTC	838

DB	847	TCAGTTTCATACCGGTTGGAACACTGGCAGACAATTTGTGATTTGGCGTATCTTTCTCGTTTC	906
QY	839	CTCCTTAGCAAGATATCTTCAGCAAAAACGCAAGGTTTTTTTGGTGTACAGCAATG	898
DB	907	CTTCTGTTTGGAAAGTACATCGAAAGAAAGAAAGGAGCTTTTTCTGGTGCCAGCTATT	966
QY	899	CGCCATTTGACGTCGTTTATATTTGGGAAGTCTCTTTGGTTTATTTTCACTCAAGCGAGAA	958
DB	967	GCTCCTATATTTTCAAGTATTTAGTCAACATTTTGTATATATTTCTCTGTTGGGCAAG	1026
QY	959	CACGGTGTGAAGTGAATAGAGAACTGAAGAAAGGTTTGAATCCACCATCTACAAAT	1018
DB	1027	CAAGGAGTTTCAATAGTGAAGCAATTGAACAGGAAATCAACCCATCATCATGATACAA	1086
QY	1019	CTGGTATTTGTCGCTTACATGACTACAGCTCTCAAACTGGCATTGCTGTTGGGATC	1078
DB	1087	ATTTATTTTCAACCGGCCCCATTTGTTGCAAAAGGTTTCAAGATCGGTGTTGTTGGGCA	1146
QY	1079	ATATCACTTGGGAAGGAATAGTAGTAGGAAGAGCTTTGCAATGTATATAAAATTTACAAT	1138
DB	1147	GTGTTTGAAGAGCTGTAGCTATTGGAAGGACATTTGCTGTATGAAGGACTACAG	1206
QY	1139	ATTGATGGCAACAAAGAGATGATAGCTATTGGGACCAATGAACGATGTTGTTCTTCA	1198
DB	1207	TTAGATGAAACAAAGGAGATGTTAGCACTTGGAAACCATGAACATAGTAGGCTCAATG	1266
QY	1199	TCTTGTACTTCAACAGGACCAATTTTCGGCTGCGCTGCGCTGGAATATAACGCTGGATC	1258
DB	1267	TCCTGTATGTCAACAACAGGTTCTTCTCAGCTTGGCAGTTAACTTCTCATGGCTGGCTG	1326
QY	1259	AAGCAGCAGCTTCCAACTTATATGATGCACTTGCAGTAAATGTTGACATTTGTTCTCTG	1318
DB	1327	AAGACTCTGTATCCATGTGTTATGTCAGTAGTGGTCTTCTTCTTCTTCTTCTTCTTCT	1386
QY	1319	ACACCTTGTTCATTTACACTCCCTGGTGGTGTATCAGCTATATCGTATCTGCAATG	1378
DB	1387	ACACCGCTATTCAAATAFACACCGAATCAATCTTAGGTCGATCATTTATTTCTGCGGTG	1446
QY	1379	CTTGAATCATAGATTAAGAGCAGCCATCATCTATTTTAAGGTTGACAAATTTGACATTT	1438
DB	1447	ATCGGCTTGTGGACTACGAAGCAGCAATCTCATCTGGAAGTTGAACAATTTGACATTC	1506
QY	1439	GTGGTGTGCATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTC	1498
DB	1507	ATTGCTTGCATGGAGCTTTTTTGGTGTGTTTGTGATCCGTTGAGATTTGGCTCTTG	1566
QY	1499	ATAGCTATTGTAATCTGTACTTTGGGTACTTCTTATTTATGCAAGCCCAAGGACATTC	1558
DB	1567	ATTGCTGTAGCAATCTCAATTTGCCAAAATACTTCTTCAAGTAAACAAGGCCAAGGACGC	1626
QY	1559	GTTTTGGGCAACATTCCAAATTTGTGTATATACCGAATTTGAGCACTATCAAAATGCA	1618
DB	1627	CTACTTGGAAACCTTCCCGGCAACCATATATACCGGAAATCATCAGCCAGTATCCAGA	1686
QY	1619	AAACATGTTCTCGAATCTAATTTCTAGAGATTGATGACCAATTTACTTTTGCCAAATGCC	1678
DB	1687	AAACTTACTCTCTGGGTTGATTTGAGGTTGATTTCTGCTATTTATTTTTCCAACTCT	1746
QY	1679	AGCTATTTAAGAAAGGATCAAGGTTGGATTTGATGAAGAGAAAGAAAGATTTAAAGCT	1738
DB	1747	AAATACGTCGGAAGAAATTTCTTAGTGGCTGACAGACGAAGAAAGACAGAGCTAAAGCA	1806
QY	1739	ACAGGGGAGACTAGTTTGCAGTATGTTATATTAATGATGATGCTGTTGGAAACATTTGAT	1798
DB	1807	GTGGGATTCCTTAAATCAGTTTCTGATTTGGAAATGTCCGCGGTCATCCGACATCGAT	1866
QY	1799	ACAAGTGAATTAAGTATCTTGAAGCTGAAGAGATTTACAGAGAGAGAGAGCTACAG	1858
DB	1867	ACAAGCGGATACATGCTCTTGAAGTCTATFACAGAAATCTTCAAGAAAGATATGAG	1926
QY	1859	CTTGTGTTTGGTCAATCTCTGTAAGTGAAGTGAAGAACTGAAACAAATTCGAAGTTCCAA	1918

Db 1927 CTCATTCTGCGAATCTCGTTCGGTGCATAGAAAAAATCGAAGCGTCGAAGCTCAAC 1986  
QY 1919 AATCAATTAGGGAAGAAATGGATCTATCTGACTGTTGGAAGAGCGCGTTGGAGCATGCA 1976  
Db 1987 GAGCATTGGAAGCAGCAATATATCTCTCGCGGTCTCTGACGCTGTGCGATTCTGTA 2044

RESULT 6  
US-09-720-317A-7  
; Sequence 7, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 780  
; TYPE: DNA  
; ORGANISM: Helianthus tuberosus  
US-09-720-317A-7

Query Match 13.2%; Score 324; DB 4; Length 780;  
Best Local Similarity 70.6%; Pred. No. 1.2e-79;  
Matches 432; Conservative 0; Mismatches 180; Indels 0; Gaps 0;  
QY 1400 GCAGCCATCCATCTATTAAAGTTGACAAATTTGACTTGTGTGTCATGATGTCATAC 1459  
Db 17 GCCGCGATTCCACTCTGGACACTAGACAAATTCGACTTGTGTATGCATGATGTCATAC 76  
QY 1460 ATTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCTATTGTAAATATCTGTA 1519  
Db 77 TTGTGTGTGTCTTTGGAGTGTGAAATTTGGATTAGTTATCCGGTGCATTTGCTGTG 136  
QY 1520 CTTTGGGTACTTCTATTATTTGAAGCCCAAGACATTCGTTTGGGCAACATTCCAAAT 1579  
Db 137 CTTAGGGTACTCTATTGTCTCGAGGCCAAGAACATCGACGCTAGGTCTCATACCGCAT 196  
QY 1580 TCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACATGTTCTGGAATGCTA 1639  
Db 197 TCCACTATTATAGAAGTATGGATCAATACCAAAATGCGAAAAGCGTTCCAGGAATCTTG 256  
QY 1640 ATTCTAGAGATTGATGCACCAATTTTACTTTGGCAATGCCAGCTATTTAAGAGAAAGGATC 1699  
Db 257 ATACTTCAAATCGAAGCACCATTATTTACTTTGCTACTCTAGCTATTGAGGGAAGGATT 316  
QY 1700 ACAAGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACAGGGGAGACTAGTTTGCAG 1759  
Db 317 GTGAGATGGGTTTCAAGAGGAAGATAGGTTGAAGTCTTTAAAGGAGAATGACTTTGCAA 376  
QY 1760 TATGTTATATTCATATGATGCTCTTTGNAACATTTGATACAAAGTGAATAGTATGCTT 1819  
Db 377 TATGTCATTTCTGATTTGATGCTGTTTGGAAATATTGATACAAAGTGGGATAAATGCTT 436  
QY 1820 GAAGAGGTGAAGAAGATTACAGAGAGAAGAGACTACAGCTCTGTTTGTGTCATCTGTA 1879  
Db 437 GGAGAAGTTAAAGGTTATGGAAGAGAGAGGCTAAAGTTGTTTTAGCGAATCCGGGC 496  
QY 1880 AGTGAAGTGAAGAAGAACTGAACAAATCGAAGTTCCAAAATCATTTAGGGAAGAAATGG 1939  
Db 497 GGAGAGTAATAAGAAGATGAACAAAGCGAAGTTGATAGAGGTGATCGGCAAGATGG 556  
QY 1940 ATCTATCTGACTTTGAAGAGCGCGTTGGAGCATGCAACTTCAATCTAGTGAAGCAAAA 1999  
Db 557 ATATATCTAACAGTGGGAGAGCGGTTGGAGCGTGCAACTTTATGCTTCTATCTTACAAG 616

QY 2000 ACGAACCCCAAG 2011  
Db 617 AACGCCGAAG 628  
RESULT 7  
US-09-720-317A-5  
; Sequence 5, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 1240  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-720-317A-5  
Query Match 8.8%; Score 216.2; DB 4; Length 1240;  
Best Local Similarity 53.4%; Pred. No. 1.4e-49;  
Matches 477; Conservative 0; Mismatches 413; Indels 3; Gaps 1;  
QY 1085 CTTGGGGAAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATAAAAAATTACAATATTGAT 1144  
Db 9 CTCACGGAAGCTATGCCGTTGGCCGATCTTTTCGCTCCGTAAGAGGTTACAGACTCGAC 68  
QY 1145 GGCACAAAGAGATGATAGCTATTGGGACCATGAACGTAAGTGTGTTCTTTCACTTTC 1204  
Db 69 GGCACAAAGAGATGCTGGCCATGGGTTCTCCAACGTTGCTGTTCTCTGCTCTGTCG 128  
QY 1205 TACCTCACAACAGGACCAATTTTCGGTTCGGCTGTGAACTATAACGCTGATGCAAGACA 1264  
Db 129 TATGTGGCAACAGGTTGCTTCTCCGAAACGGCAGTGAACCTTCAGCGCGGGGCGAGTCG 188  
QY 1265 GCAGCTTCCAACTATATATGTCACCTTGCAAGTAATGTTGACATTTGTTATTCCTGACACC 1324  
Db 189 ACGTTTCAAACATCGTCATGTCATCACCGTTCCTGTCACCTGGAGCTGTTCTAGAAG 248  
QY 1325 TTGTTCCATTTACACTCCCTGGTGGTGTATCAGCTATPATCGTATCTGCAATGTTGGA 1384  
Db 249 CTCCTCTACTACACGCCCATGGCGTCTCGCTCCATCATCTCTGCGGCTCTTCCGGGA 308  
QY 1385 CTCATAGATTATGAAGCAGCCATCCATCTATTTAAGTTGACAAATTTGACTTTGGTG 1444  
Db 309 CTGATCGACATCAAGAGGCGCTTCAGACATATGGAAGATCGACAAGATGATTTCTCACC 368  
QY 1445 TGCATGAGTGACATACATTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCT 1504  
Db 369 TGCTCTGGTGGCTTTGTTGGCGTCTGTTGGGTGGTGGAGATTTGGGCTTCGAGTTGCA 428  
QY 1505 ATTGTAATATCTGTAATCTGGGTACTTCTATTTATTTGAAAGGCCAAGGACATTCGTTT 1564  
Db 429 CTTGGCATTTCTCTCGAAAGATCATACAGTCGCTTCGGCCTCAGGTGGAGATCTT 488  
QY 1565 GGCACATTCCAAATCTGTGATATACCGAATGTTGACACTATCAAAATGCAAAACAT 1624  
Db 489 GCGAGGCTACAAGGGACAGATATCTTCTGACGCTCAGGCACTACCTGTAGCCTGCTTA 548  
QY 1625 GTTCTCGGAATGCTAATCTAGAGATTGATGACCAATTT---TACTTTCCCAATGCCAGC 1681  
Db 549 ACTCGACTGTACTGCTTATCGGCTGACACATCTCTCTCTGCTTCTATCAACGCCACT 608  
QY 1682 TATTTAAGAGAAAGGATCAAGAGTGGATTGATGAAGAGAAAGAAATTAAGACTACA 1741

Db 609 TCCGTCAAAGAAAGGATCACAGATGGGTTTGGGAAGGAGTGGAGACCTCAAATGGAAAA 668  
Qy 1742 GGGGAGACTAGTTTGCAGTATGTTAATTTGATATGAGTGCCTTTGGGAACATTGATACA 1801  
Db 669 GCGAGGGAGAGGATACAAGCAGTTGTCTTGATATGTCAAGTGTGGTAAACATCGACACT 728  
Qy 1802 AGTGGAAATAGTATGCTTCAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTT 1861  
Db 729 TCAGGACTCAGCTGCACTGGAAGAAATACACAAGGAGTTGGTGTCTCTTGGCTTACAGATG 788  
Qy 1862 GTTTTGGTCAATCTGTAAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCAAAAT 1921  
Db 789 GCTATAGCCAGTCGCGGATGGAAGGCAGTTCAGAGATGAAAGTGTCAAGGTGTGGAC 848  
Qy 1922 CATTTAGGAAGAAATGGATCTATCTGACTGTTTGAAGAGCCGTGGAGCATG 1974  
Db 849 AGGGTAGGACGAGCTGGATCTTCATGACAGTAGTGAAGCGGTGGAGCCGTG 901

RESULT 8

US-09-720-317A-21  
; Sequence 21, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 21  
; LENGTH: 2022  
; TYPE: DNA  
; ORGANISM: *Triticum aestivum*  
US-09-720-317A-21

Query Match 8.4%; Score 205.6; DB 4; Length 2022;  
Best Local Similarity 47.0%; Pred. No. 1.7e-46;  
Matches 791; Conservative 0; Mismatches 869; Indels 24; Gaps 4;  
Qy 306 AGTCTTGAAGCTGACCTCATAGCTGGCATCACCATCCGTAGCTTGGCCATTCCTCAGG 365  
Db 103 AGGACTTCCAGGCCGACCTCGCGCGGCATCACTGTGCGCGTCATGCTTGTGCGCTCAGG 162  
Qy 366 GCATCAAGTTATGCCAAGCTCGCCAACTCCCTCCAAATTTCTTGGACTATATTTCGAGCTTTA 425  
Db 163 CAATGTCTATGCAAGAGCTGGCTGGGCTTTCACCCAAATTTATGGGCTCTACACAGGCTTG 222  
Qy 426 TACCACATTTGATTTATGGCATGATGGTAGCTCGAGGATTTGGCAGTGGGACTGTGG 485  
Db 223 TCCACTATTGCTACGCAATTTTGGTCCCTCAGCAATGACAGATGAGTCCAGTGG 282  
Qy 486 CGGTGGATCGCTTCTGATGGTTCGATGTTGAGTAATCCCGTTGATCCCAATGAAGACC 545  
Db 283 CACTTGTCTCTGCTAGTGTCCAAATGTTCTTGGGGGTATAGTTAATTCATCTAGTG--- 339  
Qy 546 CAAAGCTTTACCTCCAGCTGGCTTTTCACAGCTACATATTTCGTGGTGTGTTTTTCAGGCTG 605  
Db 340 ---AGCTGTACACGGAATTAGCCATATTATTGGCAATTCATGTTGGAACTAGGAATGCT 396  
Qy 606 CTTTGGTCTGTTTGTAGTTGGGTTGATCGTGATTTTCTGTCACATGCAACCATAAATAG 665  
Db 397 TGAAGCAATGCTAAGACTTGGCTGGCTTATTTCGTTTCATAGCCATTCGTAAATATCTG 456  
Qy 666 GGTTCATGGAGGAGCAGCCAGCTGGTGTGCTGCAGCAACTTAAATCGATTCCTTGGCC 725  
Db 457 GATTCACTACAGCTTCGGCCATCGTAATTGGTTGTTCGCCAAATCAAGTATTTCTTGGG-- 514

Qy 726 TTGACATTTTCAACCATGGAGCTGATATCATATCAGTGAATCGGCTCTGTGTTTTCACCCAAA 785  
Db 515 -TTACAGTGTACAAAGTAGCAAAATTAACCACTTATTGAGAGTATAAATTCGTGAA 573  
Qy 786 CTCATGATGGAGGTGGGAAAGTGTGTGTTAGAGATGTCTTCATTTCTTCTCCCTCTTA 845  
Db 574 TAGATCAGTCTCTCGCTGCCCTCCATTGTTAAATGGGATCAGCGTTTCTGTTATTCTTCTAA 633  
Qy 846 GCACAAGATATCTCAGCAAAAACGACCAAGTGTGTTTGGGTGTCACCAATGGGCCAT 905  
Db 634 TAAAGAAAAGCTAGGAAAACAAATAAAATTTACGTTTCTTGAGAGCTTCTGTGTCAC 693  
Qy 906 TGACGTCCGTTTATATTTGGGAAAGTCTCTTTGGTATTATTTCACTCAGCCCGAAGCAGCGTG 965  
Db 694 TAAAGCTGTGTTCTTGGACATTTGTTGTGAAAAATTTTCCGTCC-----AACGTCCA 747  
Qy 966 TTGAAGTATAGGAGAACTGAAAGAGGTTTGAATCCACATCACTCAAAATCTGGTAT 1025  
Db 748 TATCAGTGGTAGTGAAATACCGCAAGGCTTCCAGGTTTCTCCATTCTCGAGGATTG 807  
Qy 1026 TTGTGTCGCTTACATGACTACAGCTGTCAAACTGGCATTTGCTTGGCATCATATCAC 1085  
Db 808 AACATCTGATGTCCCTAATGCGCAACTGCAATCTTATCAGTGGTTG-----CTA 858  
Qy 1086 TTGCGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTTACAATATTGATG 1145  
Db 859 TTTTGGAGTCTGTGGGATTGCTAAAGCGTTAGCTGCGAAGAATGGTTATGAGTTGGACT 918  
Qy 1146 GCACAAAAGAGATGATAGCTATTTGGGACATGAAGTAGTGTGTTCTTTCACTCTTGT 1205  
Db 919 CAAAACAAAGAGTTATTGGCCCTTGGCTTATCAAAATATATGGCGTTTCATTTCTCTCGAT 978  
Qy 1206 ACCTCACAACAGGACCATTTTCGGCTTCGGCTGTGAATATAAAGCTGATGCAAGACAG 1265  
Db 979 ATCTGCTACAGGCTCTCTTTCTAGGCTGCTGTGAATCATGAAGCGGGGAAAGACTG 1038  
Qy 1266 CAGCTTCCAAACATTTAATGTCACTTGCAGTAATGTTGACATTTGTTATTCCTGACACCT 1325  
Db 1039 GATTTATCAGGAATCATATGCGCATATAATTTGCAAGTCTCTCTTTTATGACACCAT 1098  
Qy 1326 TGTTCATTTACACTCCCTGGTGGTCTATCAGCTATTATCGTATCTGTAATGCTTGGAC 1385  
Db 1099 TATTACTGATATACCTCAGTGTGCATTTGGCTGCCATTGTGATTTCTGCTGCTGCTG 1158  
Qy 1386 TCATAGATTTCAAGCAGCCATCCATCTATTAAAGGTTGACAAAATTTGACATTTGTTG 1445  
Db 1159 TGGTAGATTGAAGAGGCCATCTTCTGTTGGGGTATTGATAAGAGGATTTCTTTCTGT 1218  
Qy 1446 GCATCAGTGCATACATTTGGCGTGTCTTTTGGCAGTGTGAAATTTGGCTTTAGTCA 1505  
Db 1219 GGGCGATGACATTTTACTACAACCTTAACTTTTGGCAATTTGAGATTTGTTGTTGGG 1278  
Qy 1506 TTGTAATATCTGACTTCCGGTACTTCTATTATTATTGCAAGGCCAAGGACATTCGTTTGG 1565  
Db 1279 TCGGGTCTTTCGTGGCAATTTGTGATCCATGAATCTGCAAAATCCGCATATAGCTGTTTGG 1338  
Qy 1566 GCACATTTCCAAATTTCTGTGATATACCGAAATTTGAGCACTATCAAAATGCAAAACATG 1625  
Db 1339 GCGGTTTCCCTGGCCACCACTGTGTACAGGAATACATTCGAGTACCTCAGGCTTATACAT 1398  
Qy 1626 TTCTCGAATGCTAATTTCTAGAGATTGATGCACCAATTTACTTTTGCCAAATGCGAGTAT 1685  
Db 1399 ACACGGGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1458  
Qy 1686 TAAGAGAAGGATCAAGAGTGGATTGATGAAGAGAAGAAAGAAATTTAAAGCTACAGGG 1745  
Db 1459 TAAAGGACAGGTTGCGTGAATGAGCTCAAACTCCCAAAATTTCAAAACCGTGGACCTGATG 1518  
Qy 1746 AGACTAGTTTGCAGTATGTTAATTTGATATGAGTGTGTTGGAACATTTGATACAAGTG 1805  
Db 1519 TTGGAAGGAGTGTACTTTGTGATCTCCGAGATGCCCTGTTTACATACATCTGACTCGAGG 1578  
Qy 1806 GAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTTACAGCTTGT 1865

Db 1579 CTGTTCAAGCTCTCAAGGACCTGCACCAAGAAATACAAAGCAGCGACATCCAGATTGCTA 1638  
Qy 1866 TGTCTAATCTGTAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCCAAAATCATT 1925  
Db 1639 TAGCGAATCTTACCGCGAGTGACACTATTGCTGTCAAGAGCGGCATCATGACATGA 1698  
Qy 1926 TAGGGAAGAAATGATCTATCTGACTGTGTAAGAGCGCGTTGAGCATGCAAACTTCAATC 1985  
Db 1699 TTGGCGCAGGTTGTTGCTCGAGTGACACGCGGTGCAAGTATGCTCCAGCATG 1758  
Qy 1986 TACG 1989  
Db 1759 TGCG 1762

RESULT 9  
US-09-720-317A-13  
; Sequence 13, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 13  
; LENGTH: 493  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-09-720-317A-13

Query Match 4.8%; Score 118; DB 4; Length 493;  
Best Local Similarity 55.0%; Pred. No. 1.7e-22;  
Matches 252; Conservative 0; Mismatches 205; Indels 1; Gaps 1;  
Qy 816 TAGGATGTGTCTTCATTTCTTCTCTTAGCAACAAGATACTTCAGCAAAAAACGACCAA 875  
Db 36 TCGATGCTCTTCTCTATATTCATCTCACCACACGGTTTCATCGGAGAGGTACAAGA 95  
Qy 876 GGTGTTTTGGGTGTCAGCAATGCGGCATTTGACGTCGCTTATATTGGGAAGTCTCTTGG 935  
Db 96 AGCTGTTCTGGCTGTGAGCGATCTCGCCTCTGCTGTGCTCTCTCTGTCACCGCTGCGG 155  
Qy 936 TTTATTTCACTACGCGGAGACGAGTGTGTAAGTATAGGAGAACTGAAGAGGTT 995  
Db 156 TCTACGCGAACGGGTGACAGCGACGCGGTCAAGATCATCCAGAAAGGTGCAAGCGGCC 215  
Qy 996 TGAATCCACCATCACT-CACAAATCTGGTATTGTTGTCGCTTATACATGACTACAGCTGC 1054  
Db 216 TAAACCAAGCTCCGTGGAAGAGATACCTCAACGGCGCGACACAGGAGTGGCGC 275  
Qy 1055 AAAAATGGCAATGCTGTGGCATCATATCACTTTCGGAAGGAATAGCAGTAGGAAGAC 1114  
Db 276 CAAGATCGCGTTCATCTGGCATCATCGCCTCACGGAAGTATCGCGTGTGCGCATCT 335  
Qy 1115 TTTGCAATGTATAAAAATATACATATTGATGGCAACAAGAGATGATAGTATTGGAC 1174  
Db 336 TTGCGCTCCGTAAGAGGGTACAGACTCGACGGCAACAAGAGATGCTGGCCATGGGGTTC 395  
Qy 1175 ATGAAGTATGTTGTTCTTCTGCTCTGCTACTTCACAAACAGGACATTTTTCGCGTTCG 1234  
Db 396 TCCAAAGTTGCTGGTTCTCTGCTCTGCTGCTATGCTGGCAACAGGTTCTTCTCCGAA 455  
Qy 1235 GCTGTGAATATTAACGCTGGATGCAAGACAGCAGCTTC 1272  
Db 456 GCAGTGAACCTTCAGCGCGCGGGGCCAGTCGACCGTTTC 493

RESULT 10  
US-09-720-317A-9  
; Sequence 9, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 9  
; LENGTH: 484  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (398)  
; OTHER INFORMATION: n = A, C, G or T  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (483)  
; OTHER INFORMATION: n = A, C, G or T  
US-09-720-317A-9

Query Match 4.8%; Score 117.4; DB 4; Length 484;  
Best Local Similarity 55.7%; Pred. No. 2.5e-22;  
Matches 265; Conservative 0; Mismatches 207; Indels 4; Gaps 2;  
Qy 717 TTCTGGCCTTGAGCATTTCCACCATGGAGCTGATATCATATCAGTGATGCCCTCTGTTT 776  
Db 2 TTACAGGCATCAAGAGCTTTACAAAGAAAACCGATATAATTTCCGTGATGAGCTGAAGTC 61  
Qy 777 TCACCAAACTCATGA---GTGGAGGTGGGAAAGTGTGTGTAGTAGTGTCTTTCATT 833  
Db 62 CAACACAGAGCGCATAAACAGGTGGAAATTTGGCAAACTATTGTGATTGGCATAACTTTCC 121  
Qy 834 TCTTCTCTCTTAGCAACAAGATACTTCAGCAAAAAACGCAAGGTTTTTTTGGGTGTCAG 893  
Db 122 CATTCTCTTCTGCTTCCCAAGTACATTGGAAAGAGAAATAGGAAGTTCTTCTGGTGCCAG 181  
Qy 894 CAATGGCGCCATTGACGTCCTGTTATATTGGGAAGTCTCTTGGTTTTTATTTCACTCAGCG 953  
Db 182 CTATTGCTCTTATACTTTCAGTTATTTTGGCAACCTTTTGTGTTCTTACTCTGTCGTCG 241  
Qy 954 AGAAGCAGCGTGTGAAAGTGAAGAGTGAAGAACTGAAGAGGGTTTGAATCCACCATCTCA 1013  
Db 242 ACAAGCAAGGTGTTCCAGATTGTTAAACCATCAAAAAGGGCAATAAACCCATCATCAGTCC 301  
Qy 1014 CAAATCTGGTATTGTCGCTTACATGACTGCTGCAAGTGTGCAAACTGCAATGTCGTTG 1073  
Db 302 ACAAAATTTATTTCATGCTGCTCAATTTTGGAAAAGGTTTCAAGATCGGTGTCATTCCG 361  
Qy 1074 GCATCATATCACTTCGGGAAGGAATAGCAGTAGGAAGAGCTTTTGCATGTATAAAAT 1133  
Db 362 CATGATCGTTTAAAGGAGCTGTGGCAATTTGGGNGACGTTTGTCTGCTCTGAAGACT 421  
Qy 1134 ACAATATTGATGGCAACAAAGAGATGATGATTGGGACCAATGAACGATGATGTTGGT 1189  
Db 422 ATCAATATTGATGGCAACAGGAGATG-GAAACATTGGAACATATGAACATACAAGGT 476

RESULT 11  
US-09-785-381-4  
; Sequence 4, Application US/09785381  
; Patent No. 6602992



```
; GENERAL INFORMATION:
; APPLICANT: DALLOS, Peter
; APPLICANT: ZHENG, Jing
; APPLICANT: NADISON, Laird
; TITLE OF INVENTION: A MAMMALIAN PRESTIN
; FILE REFERENCE: 0290-3701
; CURRENT APPLICATION NUMBER: US/09/785,381
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,461
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2441
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-785-381-4

Query Match      3.7%; Score 90.4; DB 4; Length 2441;
Best Local Similarity 46.9%; Pred. No. 2e-14;
Matches 317; Conservative 0; Mismatches 356; Indels 3; Gaps 1;

QY 1061 GGCATTGTGCTGGCATCATATCACTTGGGAAGGAATAGCAGTAGGAAGAGCTTTGCA 1120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1233 GCCATTGGCATGCCATCGTTGGATTTTCAGTGACGATCTCCATGGCCAAACCTTTGGCA 1292
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1121 ATGTATPAAATAATPACAATATTGATGGCAACAAAGAGATGATAGCTATTGGGACCATGAAC 1180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1293 AATAAGCATGGCTACCAGGTTGATGGCAATCAGGAGCTCATTTGCCCTTGGGGATATGCAAC 1352
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1181 GTAGTTGGTCTTTACCTCTTGCTACCTCAACAGGACCAATTTTCGGGTTCCGGCTGTG 1240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1353 TCCATTGGATCTCTCTTCCAAACCTTCTCGATTTCTGCTCTGTCGTCGAAAGCCTTGT 1412
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1241 AACTATAACGCTGGATGCAAGACAGCAGCTTCCAACATTATATGTCACCTTGCAGTAATG 1300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1413 CAGGAAGGAACCTGGAGGGAACACAGCTTGCAGGTTGTTGGCCCTGTTGATGATCTG 1472
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1301 TTGACATTTGTAATTCCTGACACCCCTTGTTCATTAACATCCCTGGTGGTGTATCAGCT 1360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1473 TTGGTCATATTAGCCACCGGATTCCTCTTTGAGTCGTATACCCAGGCTGTCTTTCCGCC 1532
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1361 ATTATCGTATCGCAATGCTTGGACTCATAGATTGAAGCAGCCATCCATCTATTATTAAG 1420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1533 ATTGTGATCGTCAACCTGAAAGGAATGTTTCATGCAAGTTCTCAGACCTGCTCTTTTGTG 1592
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1421 GTTGA---CAAAATTCACATTTGTGGTGTGATGAGTGCAATGCGGTGCTTTTGGC 1477
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1593 AGAACAGCAAAATAGAGCTGACCATCTGGCTGACCACTTTGTGTCTCCCTGTTCTC 1652
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1478 AGTGTGAAATTTGGCTTAGTCATAGCTATTGTAATACTGTACTTCGGGTACTTCTATT 1537
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1653 GGCTTGGACTACGACTGATTACCGCGTGATCATTTGCTCTGCTCAGTGATTTATAGA 1712
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1538 ATTGCAAGGCAAGGACATTCGTTTGGGCAACATTCCAAATTCGTGTATATACCGAAT 1597
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1713 ACACAGAGTCCAAAGCTACAAAGTCTGGGGCAGCTCCCTGACACGGATGTGACATTGAC 1772
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1598 GTTGAGCACTATCAAAATGCAAAACATGTTCTTGGAAATGCTAATTTCTAGAGATTGATGA 1657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1773 ATAGATTGCATATGAGGAGGTGAAGAAATTCCTTGGAAATTAATAATATCCAAATAAATGCC 1832
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1658 CCAATTTACTTTGCCAATGCCAGCTATTTAAGAGAAAGGATCAACAGGTGGATTGATGAA 1717
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1833 CCAATTTACTATGCAATAGCGACTTGTATAGCAGCGCTTTTAAAGAAAGAAAGCTGAGTA 1892
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1718 GAAGAAGAAAGAAATTA 1733
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1893 AACCCAGCACTCATTA 1908
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

; Sequence 5627, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5627
; LENGTH: 4927
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5627

Query Match      3.4%; Score 82.6; DB 4; Length 4927;
Best Local Similarity 47.0%; Pred. No. 4.3e-12;
Matches 290; Conservative 0; Mismatches 324; Indels 3; Gaps 1;

QY 1087 TCGGAAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAATTAACAATATTGATGG 1146
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1307 TGCTATTGCAAGTGTGATGAGGAAAGTATATGCCACCAAGTATGATTACACCATCGATGG 1366
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1147 CAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTCTTTTCACTCTTGCTA 1206
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1367 GAACAGGAATTCATTGCTTTGGGATCAGCAACATCTTCTCAGGATTTCTCTCTTGT 1426
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1207 CCTCAACAGGACCATTTTCGCTTTCGCTGTGAACATATAACGCTGGAATGCAAGCAGC 1266
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1427 TGTGGCCACCACTGCTCTTTCCGACGCGCTCCAGAGACATGAGGAAACACACA 1486
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1267 AGCTTCCAAACATTAATAATGTCACTTGCAGTAATGTGACATTTGTTATTCTCGACACCT 1326
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1487 GGTGCTGGCATCATCTCTGCTGCGATTTGTGATGATCGCAATTTCTGCGTGGGAAGCT 1546
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1327 GTTCCATTACACTCCCTCGTGGTGGTCTATCAGCTATTATCGTATCTGCAATGCTTGGACT 1386
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1547 TCTGGAACCTTTGCAGAAAGTCGGTCTTGGCAGCTGTTGTAATTCGCAACCTGAAAGGGAT 1606
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1387 CATAGATTATGAAGCAGCCAT---CCATCTATTAAAGTTGACAAATTTGACTTTGTG 1443
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1607 GTTTATGAGCTGTGTGACATTCCTCGTCTGTGGAGACAGATAAGATTGATGCTGTTAT 1666
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1444 GTGCAATGATGATACATTTGGCGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGC 1503
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1667 CTGGGTGTTTACGTGTATAGTGTCCATCATTTCTGGGCTGGATCTCGGTTTACTAGCTGG 1726
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1504 TATTGTATATCTGTACTTCTGGGTACTTCTATTATTTCAGAGGCCAAGACATTCGTTTT 1563
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1727 CCTTATATTGAGCTGTTGACTGTGGTCTCTGAGAGTTTCAGTTTCTTCTTGAATGGCT 1786
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1564 GGGCAACATTCCAAATTTCTGTATATACGAAATGTTGAGCACTATCAAAATGCAAAACA 1623
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1787 TGAAGACATCCCTAGCACAGATATCTAANAAGTACCAAGATTAACAAAACATTTGAAGA 1846
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1624 TGTCTCTGGAATGCTAAATTTCTAGAGATTGATGACCAAAATTTACTTTGCCAATGCCAGTA 1683
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1847 ACCTCAAGGAGTGAAGATTTCTTAGATTTTCCAGTCTTATTTTCTATGGAATGTCGATGG 1906
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

; TTTAAGAGAAAGGATCA 1700
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; TTTTAAAAAATGTATCA 1923
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-09-949-016-5627
```



US-09-785-381-2  
; Sequence 2, Application US/09785381  
; Patent No. 6602992  
; GENERAL INFORMATION:  
; APPLICANT: DALLOS, Peter  
; APPLICANT: ZHENG, Jing  
; APPLICANT: MADISON, Laird  
; TITLE OF INVENTION: A MAMMALIAN PRESTIN  
; FILE REFERENCE: 0290-3701  
; CURRENT APPLICATION NUMBER: US/09/785,381  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 60/183,461  
; PRIOR FILING DATE: 2000-02-18  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 4113  
; TYPE: DNA  
; ORGANISM: Meriones unguiculatus  
US-09-785-381-2

Query Match 3.3%; Score 79.8; DB 4; Length 4113;  
Best Local Similarity 46.1%; Pred. No. 2.3e-11;  
Matches 304; Conservative 0; Mismatches 352; Indels 3; Gaps 1;  
  
QY 1060 TGGCATTGTCGTTGGCATCATATCACTTGGGAAGCAATAGCAGTAGGAAGAGCTTTGCG 1119  
DB 1249 TGCATTGGCCATAGCCATCGTTGGATTTCAGTGACAAATTCATGCCCAAAACCTTGGC 1308  
  
QY 1120 AATGTATAAAATTAACAATATTGATGGCAACAAAGAGATGATAGCTATTGGGACCAATGAA 1179  
DB 1309 GAATAAGCATGGCTACCAAGTTGATGGCAATCAGGAGCTCATGCCTTGGGGATATGCAA 1368  
  
QY 1180 CGTAGTGGTTCCTTACCTTCCTTGCCTACCTCAACAGGACCAATTTTCGGGTTGCGCTGT 1239  
DB 1369 CTCCTATCGGATCTCTCTCCAGACCTTCTCCATTCTCTGCTCTGCTCGCAGCCCTTGT 1428  
  
QY 1240 GAACATAACGCTGGATGCAAGACAGCAGCTTCCAACATTAATAATGCTCACTTGCAGTAAT 1299  
DB 1429 TCAGGAGGNACTGGAGGNAACACAGCTCGCAGGTGCTTGGCCCTCGCTGATGATCT 1488  
  
QY 1300 GTTGACATTTGTTATTCCTGACACCTTGTTCATTAACACTCCCTGGTGGTGTCTATCAGC 1359  
DB 1489 GCTGGTCATTTTAGCCACTGGAATTCCTTTGAGTCAATTTGCCCAAGGCTGTCTCGGC 1548  
  
QY 1360 TATTATCGTATCGCAATGCTTGGACTCATAGATTATGAGCGCCATCCATCTATTAA 1419  
DB 1549 CATTGTGATCGTGAACCTGAAAGGGATGTTTATGCAAGTTCTCAGATCTGCCCTTCTCTG 1608  
  
QY 1420 GGTT---GACAAATTTGACTTTGTGTGTGTCATGATGATGATGATGATGATGATGATGATG 1476  
DB 1609 GAGAACAGCAAAATAGCTGACCAATCTGGCTTACCACTTTGTGTCTCCCTGTTCT 1668  
  
QY 1477 CAGTGTGAAATTTGGCTTAGTATAGCTATTTGTAATATCTGATCTTCGGGTACTTCTATT 1536  
DB 1669 GGGCTTGGACTACGGACTGATTAATCTGTGATCAATGCTCTGCTGACTGTGATTTACAG 1728  
  
QY 1537 TATTGCAAGCCCAAGGACATTCGTTTGGGCAACATTCGAAATCTGTGATATACGGA 1596  
DB 1729 AACCCAGAGTCCAGCTACAAGGTCCTGGGGAGCTCCCTGACACCGATGTATACATTGA 1788  
  
QY 1597 TGTGTGACCATCAAAATGCAAAACATGTTCTGGAATGCTAATTTCTAGAGATTGATGC 1656  
DB 1789 CATAGACGATATGAGGAGGTGAAGAAATTCCTGGAATAAAATATTCAGATAAAGC 1848  
  
QY 1657 ACCAAATTTACTTTGCGCAATGCCAGCTATTTAAGAGAAAGGATCACAAGGTGATGATG 1715  
DB 1849 CCCAAATTTACTATGCAACAGTCACTTGTATAGCAACGCCCTTAAAGAAAGAGCTGGT 1907

RESULT 14  
US-09-248-796A-6643  
; Sequence 6643, Application US/09248796A

; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 6643  
; LENGTH: 2487  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-6643

Query Match 3.2%; Score 77.6; DB 4; Length 2487;  
Best Local Similarity 48.8%; Pred. No. 7.2e-11;  
Matches 209; Conservative 0; Mismatches 219; Indels 0; Gaps 0;  
  
QY 257 CTTCAAGTTCTTCTCCCAATTTTCCGAATGGGCTCCCAATACACCTTTTCAGTTCTTGAAA 316  
DB 253 CTTAATTTCAATGTTTCCCAATCTTAAATGGATTTTGCATTACAATATTAGATGGTTATAT 312  
  
QY 317 GCTGACCTCATAGCTGGCATCACCATCGTAGCTTGGCCATTCCTCAGGGCATCAGTTAT 376  
DB 313 GGTGACTTGGTTGCTGGTATTAATCTGCTGGAGTTGTTTGTAGTGCCTCAATCTATGTCATAT 372  
  
QY 377 GCCAAGCTCGCCAACTCCCTCCAATCTTGGACTATATTCGAGCTTTTATACCAACCATG 436  
DB 373 GCCCAATTAGCTGGGTTAGAGCTCAATAGTTTATATCTTCTTCTGCGTGTTC 432  
  
QY 437 ATTTATGCGATGATGGGTAGCTCGAGGGATTTGGCAGTGGGACTGTGCGGTTGGATCG 496  
DB 433 ATTTATTCATTTTTCGCCACCTTCAAAAGATGTTTCCATCGTCTGCTGTATGTCC 492  
  
QY 497 CTTCTGATGGGTTCGATGTTGAGTAATGCCGTTGATCCCAATGAAGACCAAGCTTTAC 556  
DB 493 TTGCAAGTGTCTAAGGTCATTCCTCATGTTCAAGATAAGTTTGGTGACAAGTATGCTGCT 552  
  
QY 557 CTCACCTGGCTTTCACAGCTACATTTATTTGCTGTGTTTTCAGGCTGCCCTTGGGCTCG 616  
DB 553 CTTGAAATGCGACATTTTGTTCATGATTTGGTGGTGTATCGTCTTGGTATGTTGTTG 612  
  
QY 617 TTTAGGTTGGGTTGATCGTGGAATTTCTGTCAATGCAACCAATATAGGGTTTCATGGGA 676  
DB 613 TTAGCTTTAGGATTCATTTAGAAATTTATTTCTATCCAGCGGTCTATGGGTTTCATGACT 672  
  
QY 677 GGAGCAGC 684  
DB 673 GGTCTGCG 680

RESULT 15  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
; US-08-232-463-14

Query Match 3.1%; Score 76.4; DB 1; Length 7218;
Best Local Similarity 6.3%; Pred.No. 2.8e-10;
Matches 26; Conservative 235; Mismatches 151; Indels 0; Gaps 0;

QY 1672 CAATGCACGCTATTAAAGAAAGGATCAAGGTGGATTGATGAAGAAGAAAGAAAT 1731
Db ||||| || ||||| || ||| : : : : : : : : : : : : : : : : : : : :
1471 CTATGCAAGTAGTTAAAGAGATAGAGAATTGGTACRRRRRRRRRRRRRRRRRRR 1412

QY 1732 TAAAGCTACAGGGGACGACTTTGCAGTATGTATTAATTGATGAGTGCTGTGGAAA 1791
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1411 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1352

QY 1792 CATTGATACAAGTGAATTAAGTATGCTGAAGAGGTGAAGAAGATTACAGAGAGAAGA 1851
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1351 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1292

QY 1852 GTACAGCTTTGTTGGTCAATCCTGTAAGTGAAGTGAAGAAGAACTGAACAATCGAA 1911
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1291 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1232

QY 1912 GTTCCAAATCATTTAGGGAAGAAATGGATCTATCTGACTGTTGAAGAGCCGTTGAGC 1971
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1231 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1172

QY 1972 ATGCAACTTCAATCTACGTGCAAGCAAAACGAAACCCAAAGAAAGATGAAACAGAGGTTG 2031
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1171 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1112

QY 2032 GAACAATGTGTGACTGAGTCATATGCAAGAGTATTCCTAAATCAATAA 2083
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1111 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1060
```

Search completed: September 12, 2005, 18:42:04  
Job time : 429 secs



QY 840 TCCTTAGCAAGATACCTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTCAGCAATGG 899  
DB 61 TCCTCGTCAAGAATACCTTCAGCAAAAACGACCAAGGTTTTCTTGGGTGTCAGCAATGG 120  
QY 900 CGCCATTGACGTCGTTATATTGGGAAGTCTCTTGGTTTATTTCTACTCAGCGCGAGAAGC 959  
DB 121 CACCACCTGACGTCGTTATATTGGGAAGTCTGCTGCTATGTACACACGCTGAGAAGC 180  
QY 960 ACGGTTTGA--GTGATAGGAGAACTGAAGAAGGTTTTGAATCCACCATCACTCACAAA 1017  
DB 181 ATGGAGTTCAAGTGTGATAGGAGAAATGAAGAAGGTTTTGAATCCACCATCACTCACAAA 240  
QY 1018 TCTGGTATTTGTCGCTTCAATGACTACAGCTGTCAAAACTGGCAATTTGTCGTTGGCAT 1077  
DB 241 TCTGGTATTTGTCGCTTCAATGACTACAGCTGTCAAAAACAGGCAATTTGTCGTTGGCAT 300  
QY 1078 CATATCACTGGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATCTATAAAAAATTACAA 1137  
DB 301 TATATCGCTTGGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATCTATAAAAAATTACAA 360  
QY 1138 TATTGATGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTAC 1197  
DB 361 TATTGATGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGTAGTTGGTTCTTTAC 420  
QY 1198 CTCTTGCTACCTCAACAAAGGACCAATTTTCGGCTTCGGCTGTGAACATAAAGCTGGATG 1257  
DB 421 CTCTTGCTACCTCAACAAAGGACCAATTTTCGGCTTCGGCTGTGAACATAAAGCTGGATG 480  
QY 1258 CAAGACAGAGCTTCAACATATTATATGTCATTCGAGTAAGTTTGACATCTGTTATTCCT 1317  
DB 481 CAAGACAGAGCATCAACATATAATGTCATTCGAGTAAGTTTGACATCTGTTATTCCT 540  
QY 1318 GACACCTTGTTCATTAACATCCCTCGTGTGCTATCAGCTATTATCGATCTGCAAT 1377  
DB 541 AACAACATTTTCCATTAACATCCCTCGTGTGCTATCAGCCATTAATGTCATCTGCAAT 600  
QY 1378 GCTTGGACTCATAGATTATGAAGAGCCATCCATCTATTATTAAGTTGACAAATTTGACTT 1437  
DB 601 GCTTGGCTCATAGATTATGAAGAGCCATCCATCTATTATTAAGTTGACAAATTTGACTT 660  
QY 1438 TGTGGTGTGCATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGT 1497  
DB 661 TGTGGTGTGCATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGT 720  
QY 1498 CATAGCTATTGTAATATCTGACTTCGGTACTCTTCTATTATTGCAAGGCCAAGGACATT 1557  
DB 721 CATAGCTATTGCAATATCTGACTTCGGTACTCTTCTATTATTGCAAGGCCAAGGACATT 780  
QY 1558 CGTTTTGGGCAACATTTCCAAATTTCTGTATATACCGAAATGTTGAGCACTATCAAAATGC 1617  
DB 781 TGTTTTGGGAAACATTTCCAAATTTCTGTATATACAGAAATGTTGAGCACTATCCAAATGC 840  
QY 1618 ARAACATGTTCTGGAATGCTAATTTCTAGAGATTGATGACCAATTTACTTTGCCAATGC 1677  
DB 841 ARAACATGTTCTGGAATGCTAATTTCTAGAGATTGATGACCAATTTACTTTGCCAATGC 900  
QY 1678 CAGCTATTTAAGAGAAGAGATCAAAAGTGGATTGATGAAGAAGAAAGAAATTAAGC 1737  
DB 901 AAGCTATTTAAGAGAAGAGATCAAAAGTGGATTGATGAAGAAGAAATTAAGC 960  
QY 1738 TACAGGGGAGACTAGTTTGCAGTATGTTATATGATATGAGTGTGTTGGAAACATTA 1797  
DB 961 TACAGGGGAACTAGTTTGCAGTATGTTATATGATATGAGTGTGTTGCTTAACATTGA 1020  
QY 1798 TACAAGTGAATTAAGTATGCTTGAAGAGTGAAGAGATTACAGAGAGAGAGCTACA 1857  
DB 1021 TACAAGTGAATTAAGTATGCTTGAAGAGTGAAGAGATTATAGAGAAGAGAGTTACA 1080  
QY 1858 GCTTTGTTTGGTCAATCTGTAAGTGAAGTGAAGAAATCGAAATTCGAAGTTCCA 1917  
DB 1081 ACTTGTGTTTGGTCAATCTGTAAGTGAAGTGAAGAAATCGAAATTCGAAGTTCCA 1140

QY 1918 AAATCATTTTAGGGAAGAAATGGATCTATCTGACTGTTGAAGAGCGGTTGGAGCATGCAA 1977  
DB 1141 AAATCATTTTAGGGAAGAAATGGATCTATCTGACTGTTGAAGAGCGGTTGGAGCATGCAA 1200  
QY 1978 CTTCAATCTAGCTGCAAGCAAAAACGAAACCCAAAAGAAAGATGAAAACAGAGGGTTGGAACAA 2037  
DB 1201 CTTCAATCTAGCTGCAAGCAAAAACGAAACCCAAAAGAAAGATGAAAACAGAGGGTTGGAACAA 1260  
QY 2038 TGTGTGACTGATGCTATATGCCAAAAGATTTCTAAATTAACCTCAAAAAGCTTTATTCGTTTT 2097  
DB 1261 TGTGTGACTGATGCTATATGCCAAAAGATTTCTAAATTAACCTCAAAAAGCTTTATTCGTTTT 1320  
QY 2098 CGTCTTACTGATGTTTACCACCTACAAATGTCGCAATGAGAATTTCTGAAATCAGCCGGAAGA 2157  
DB 1321 CGTCTTACTGATGTTTACCACCTACAAATGTCGCAATGAGAATTTCTGAAATCAGCCGGAAGA 1380  
QY 2158 AGTTTAAAGGATAGGAAAATGAAAGATGCAAGGGTCTTCTAAATTTCTCAACTCTGCAT 2217  
DB 1381 AGTTTAAAGGATAGGAAAATGAAAGATGCAAGGGTCTTCTAAATTTCTCAACTCTGCAT 1440  
QY 2218 CTTAGTTAGAAAATCTCTATGATATAGGCTGTGAAAATTAATCTTTACGTATCATGC 2277  
DB 1441 CTTAGTTAGAAAATCTCTATGATATAGGCTGTGAAAATTAATCTTTACGTATCATGC 1500  
QY 2278 TTGATAATATTTCAAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACAC 2337  
DB 1501 TTGATAATATTTCAAGAGAAATGCTAGCAACACACTCTCAGACACACTCTTTTGAACAC 1560  
QY 2338 ATGTAAAGAGGTAAAGAAAGTGTGTTGCTAGCACTCTCTCATATTTCAATTTGTAAGTAAT 2397  
DB 1561 ATGTAAAGAGGTAAAGAAAGTGTGTTGCTAGCACTCTCTCATATTTCAATTTGTAAGTAAT 1620  
QY 2398 GCCATGAGAATTTAAAATCTCTTTGGAAAAAATAAATAAATAAATAAATAAATAAATAA 2447  
DB 1621 GCCATGAGAATTTAAAATCTCTTTGGCCTTAATAATCTATTCAAAATAGAAA 1670

## RESULT 2

US-09-938-842A-1502  
; Sequence 1502, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1502  
; LENGTH: 1977  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1502

Query Match 41.6%; Score 1018.2; DB 9; Length 1977;  
Best Local Similarity 72.3%; Pred. No. 4.1e-263;  
Matches 1364; Conservative 0; Mismatches 518; Indels 5; Gaps 3;  
QY 125 CACCAAGTCGAGGTTCCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTGAAG 184  
DB 58 CACACGGTGGAGGCTCCACAACTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAG 117  
QY 185 GAGACTTCTCCCTGATGACCCCTTTGAGGCGATTCCTTGAAGAACCAAGCCAGCTTCCAAGAAG 244

Db 118 GAAACTCTGTTCCAGACGACCCCTTTTAGACAATTTAAGAACCAAAAATGCATCAAGAAA 177  
Qy 245 TTCATGCTGGCCCTTCAGTCTCTCTCCCAATTTTCGAATGGCTCCCAATACACCTTT 304  
Db 178 TTGTGTAGGCTCAAAATCTCTCTCCGATTTTCGAATGGGCAACAGCTCAATCTC 237  
Qy 305 CAGTCTTGAAGAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAG 364  
Db 238 AAGTCTTCAAAATCAGATCTCATCGCGGAATCACCATCGTAGCTCGCCATCCCTCAG 297  
Qy 365 GGCAATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATTCCTGGCATATATTCGAGCTTT 424  
Db 298 GGCATCAGTTACGCCAAATCTGTAACTTGCCTCCCAATTCCTGGCCCTTTATTCGAGTTT 357  
Qy 425 ATACCAACCATGATTTATCGGATGATGGGTAGCTCGAGGGATTTGCGAGTGGGAGCTGG 484  
Db 358 GTACCCGCCATTGGTATACCGGGTCTAGGAGTTCAGGGAGCTTAGCGGTGGGAACGGTT 417  
Qy 485 GCGGTTGGATCGCTTCTGATGGGTTCCGATGTTGAGTAATGCCGTTGATCCCAATGAAGAC 544  
Db 418 GCGGTTGCGTCTCTGTTGACAGGTGCGATGCTGAGCAAGAAGTTGATGCTGAGAAGAT 477  
Qy 545 CCAAGCTTTACCTCCACTGGCTTTCACAGCTACATATTTGCTGGTGT-TTTCAGGC 603  
Db 478 CTTAAGCTTTACCTTCACTCTGCTTTCACCGCACTTTTTCGCGCGCTTCTCGAAGCC 537  
Qy 604 TGCTTTGGGCTCTGTTAGTTGGGTTGATCGGGATTTTCTGTCATGCAACCATAT 563  
Db 538 TCTCTTGGAAATTTTCAAGGGTTAGGGTTTCATAGTGGACTTTCTATCGCATGCAACGATG 597  
Qy 664 AGGGTTTCATGGAGGAGCAGCAGCGTGGTGTCTGTCAGCAACTAAATCGATTTCTGG 723  
Db 598 AGAATTCATGGAGGAGCAGCAGCGTGGTGTCTGTCAGCAAGCTTAAAGGATTTTCGG 657  
Qy 724 CTTTGAGCATTTTCAACCATGGAGCTCATATCATATCAGTGATGCGCTCTGTTTTCACCA 783  
Db 658 ACTTAAACATTTTCAACAGCTTCAACGATGTTATCTCTGTCTGTCGTTCCGTTTCTCCCA 717  
Qy 784 AACTCATGATGAGTGGGAAGTCTGTTAGGATGCTTTCATTTTCTCTCTCT 843  
Db 718 AACTCAGAGTGGAGATGGGAAGTGGCGTCTTGGATGGGTTTCTATTTCTTCTTCT 777  
Qy 844 TAGCACAAGATCTTCAAGCAAAAACGACCAAGGTTTTTGGGTGTCAGCAATGCGCC 903  
Db 778 CTCCACAGATTTTGAAGATCAAGAAACCAAAATTTCTTTTGGTGGCGGATGCTCC 837  
Qy 904 ATTGAGCTCGTTATATTTGGGAAGTCTCTTGGTTTATTTTCACTCAGCCGAGAACGCG 963  
Db 838 TTTGACCTCAGTGATTTCTTGGAAAGTCTCTTGGTTTACTTCACTCAGCTGAGAGACATGG 897  
Qy 964 TGTGAAGTATAGGAGAACTGAAGAAGGTTTGAATCCACCATCACTCACAATCTGTT 1023  
Db 898 TGTTCAGTATAGGGGACCTGAAGAAAGGTTTGAATCCACTCTCCGGTTCTGATCTCAT 957  
Qy 1024 ATTGTGTCGCTTACATGACTACAGCTGTCAAAACCTGGCATGTGCTTGGCATCATATC 1083  
Db 958 CTTTACTTCCCTTACATGTCACAGCTGTCAAAACTGGCCTCATCTGGCATCAATGC 1017  
Qy 1084 ACTTGGGAAGGAATAGCAGTAGGAAGAAGCTTTGCAATGTATATAAAATTTACAATATGA 1143  
Db 1018 TCTCGC-GAAGGATAGCAGTGGGAGGAGTTTTCGATGTTTCAAGAACTACAAACATAGA 1076  
Qy 1144 TGCCAAACAAGATATAGCTATTTGGGACCATGAACGATAGTTGGTTCTTTCACCTCTTG 1203  
Db 1077 CGGGAACAAGATATAGCTGTTTGGAAATGATGAACATCGTGGTTCTTTCACATCTTG 1136  
Qy 1204 CTACCTCACAACAGGACCATTTTTCGGTTCGGCTGTGAATATAAAGCTGGATGCAAGAC 1263  
Db 1137 TTACCTCACAACCGGACCATTTTCAAGGTCGGCAGTGAATACACCGGGTTGCAAGAC 1196  
Qy 1264 AGCAGCTTCCAAATTTAATGTCACTTTCAGTGAATGTTGACATGTTTATTCCTGACACC 1323

Db 1197 CGCAATGTCCAAACATAGTAGGCGGATTTGCGGTTATGTTTCAACATCTCTTCTCTCACACC 1256  
Qy 1324 CTTGTTCCCAATACATCCCTCGTGTGCTATCAGCTATTATCGTATCTGCAATGCTTGG 1383  
Db 1257 GCTTTTTCATACACACCACTCGTGTCTCTCTGCGCATCATATCCGCAATGCTCGG 1316  
Qy 1384 ACTCATAGATTAATGAAGCAGCCATCCATCTATTATTAAGGTTGACAAATTTGACTTGTGT 1443  
Db 1317 ACTCATTTGACTATCAAGCTGCCATCCATCTCTGGAAGTTGACAAGTTGCACTTCTCTCGT 1376  
Qy 1444 GTGCGATGATGTCATACATTTGGCGTGGTCTTTGGCAGTGTGGAATTTGGCTTATGATACC 1503  
Db 1377 GTGCGATGAGCGCTACGTTGGGGTGGTATTCGCGAGTGTAGAGATTTGGACTCGTCTAGC 1436  
Qy 1504 TATTGTAAATATCTGTACTTTCGGGTACTTCTATTATTTGCAAGGCCCAAGCACATTCCTTTT 1563  
Db 1437 GGTGGGATATCTATAGCAGGTTTGTGCTGTTGTGTCGAGGCCCAAACTTCGGGTGA 1496  
Qy 1564 GGGCAACATTTCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAAACA 1623  
Db 1497 GGGAAACATATCCAAACAGCATGATCTATAGGAACACTGAGCAGTAGTACCCATCATCAAGAAC 1556  
Qy 1624 TGTTCCTGGAATGCTAAATTTCTAGAGATTTGATGCCAAATTTACTTTTGGCAATGCCAGCTA 1683  
Db 1557 CGTTCCTGCTATTTCTCATCTTGGGATTTGATGCTCCCATCTACTTTGCTTAATGCCAGTTA 1616  
Qy 1684 TTTAAGAGAAAGGATCACAAGGTGATTCGATGAAGAAGAAAGAAATTTAAAGCTACAGG 1743  
Db 1617 CTTGGTGAAGAATCATTAAGGTGATTTGATGAGGAGGAGAGAGGTTTAAACATCAGG 1676  
Qy 1744 GGAGACTAGTTTGCAGTATGTTTATTAATTAATGATATGATGCTGTTTGGAAAATTTGATAACAAG 1803  
Db 1677 AGAGAGCAGCTTACAATATATTTACTCGATATGTCAGCTGTTGGTAAATATCGACAACAG 1736  
Qy 1804 TGGATATGATGCTTGAAGAGCTGAAGAAGATTACAGAGAGAGAGAGCTACAGCTTGT 1863  
Db 1737 CGGTATTAGCATGATGGTGGAAATTTAAGAAAGTCTATTGACAGGAGAGCGTTAAAGTTGT 1796  
Qy 1864 TTTGGTCAATCTGTAAGTGAAGTATGAAGAACTGAACAAATCGAAGT---TCCAAA 1920  
Db 1797 ATTGTCAATTCAAAAGGAGGCTGTTGAGGAAATTTAACAGATCCAAATTCATCGGTGA 1856  
Qy 1921 TCATTTAGGAAAGAAATGGATCTATCTGATCTGTTGAAGAGCGGTTTGGAGCATGCAACTT 1980  
Db 1857 TCATTTGGCAAGAGTGGATGTTCTTACCGTAGGAAGCAGTGGAGGCTTGTAGCTA 1916  
Qy 1981 CAATCTAGTGAAGCAAAACCAACCC 2007  
Db 1917 CATGCTTACACAGTTTAAACCCGAACC 1943

## RESULT 3

US-09-938-842A-1502  
; Sequence 1502, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SAME, AND METHODS OF USE  
; FILE REFERENCE: SRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1502

; LENGTH: 1977									
; TYPE: DNA									
; ORGANISM: Arabidopsis thaliana									
US-09-938-842A-1502									
Query Match 41.6%; Score 1018.2; DB 11; Length 1977;									
Best Local Similarity 72.3%; Pred. No. 4.1e-263;									
Matches 1364; Conservative 0; Mismatches 518; Indels 5; Gaps 3;									
Qy	125	CACCAAGTCGAGGTTCCACCGCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTTGAAG	184						
Db	58	CACACGGTGGAGGCTCCACAACTCAACGTTCTTGAAGTCACTTCACTCACTCACTGAAG	117						
Qy	185	GAGACTTCTTCCCTGATGACCCCTTTGAGCGAGTTCAAGAACAAAGCAGCTTCCAAGAG	244						
Db	118	GAACCTCTGTTTCCAGACGACCCCTTTAGACAAATTAAGAACCAAAATGTCATCAAGAAA	177						
Qy	245	TTTCATGCTTGGCCTTCAGTTCTTCTCCCAATTTTCGAATGGCTCCCAATACACCTTT	304						
Db	178	TTTGTGTTAGGCTCAAAATCTTCTCCCGATTTTTCGAATGGCCACCGCTACAAATCTC	237						
Qy	305	CAGTCTTCAAAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCCTCAG	364						
Db	238	AGTCTTCAAAATCAGATCTCATCGCGGAATCACCAATCGCTAGTCTCGCCATCCCTCAG	297						
Qy	365	GGCATCAGTTATGCCAAGCTCGCCAACTCTCCCAATTCCTTGGACTATATTCGAGCTTT	424						
Db	298	GGCATCAGTTATGCCAAGCTTCTTAACCTTGGCCCAATTTCTTGGCTTTTATTCGAGTTT	357						
Qy	425	ATACCAACCAATTCATATCGGATGATGGTAGCTCGAGGATTTGGCAGTGGGGCTGTG	484						
Db	358	GTACCGCCATTTGGTATACCGGTTAGCGGTTCAAGGGACTTAGCGGTGGGAACGGTT	417						
Qy	485	CGCGTTGGATCGCTTCTGATGGGTTTCGATGTTGAGTAATGCCGTTGATCCCAATGAAGAC	544						
Db	418	GGGTTGGCTCTCTGTTGACAGTGGGATGCTGAGCAAGAAGTTGATGCTCAGAAAGAT	477						
Qy	545	CCAAAGCTTTACCTCACCTGGCTTTCACAGCTACATTAATTTGCTGGTGT- TTTCAGGC	603						
Db	478	CCTAAGCTTTTACCTTCACTTGTCTTTCACCGCACTTTTTCGCGCGGCTTCTCGAAGCC	537						
Qy	604	TGCTTGGCTCTGTTAGGTTGGGTTGATCGTGAATTTCTTCATGCAATGAACCATAT	663						
Db	538	TCTTTTGGAAATTTTCAAGGTTAGGTTTCAATGAGACTTTCATCGCATGCAACGATAGT	597						
Qy	664	AGGTTTCATGGGAGGACGCCACGCTGGTGTCTGTCAGCAACTAAATTCGATTTTGG	723						
Db	598	AGGATTCATGGGAGGAGCAGCGCTGGTGTGATCTGCAACAGCTTAAGGGTATTTTCGG	657						
Qy	724	CCTTGAGCATTTTCAACCCATGGAGCTGATATCATATCAGTGTGCGCTCTGTTTTCACCCA	783						
Db	658	ACTTAAACATTTTACAGACTCTACCGATGTTATCTCTGTCATGCGTTCGCTTTCTCCCA	717						
Qy	784	AACTCATGATGGAGTGGGAAGTCTCTGTTAGGATGCTTTCATTTTCTTCTCTCT	843						
Db	718	AACTCAAGTGGAGTGGGAAGTGGCGTTCTTGGATGTTTCTTGAATGTTTCTTCTTCT	777						
Qy	844	TAGCAACAAGATCTTCAGCAAAAACGACCAAGGTTTTTTTGGGTCTGACCAATGGCGCC	903						
Db	778	CTCCACCATATTTTACGATCAAGAAACCAAAATTTCTTTGGTGGCGCGATGGCTCC	837						
Qy	904	ATTGACGTCGGTTATATTTGGGAAGTCTCTTGGTTTATTTTCACTACGCGGAGAGCAGCG	963						
Db	838	TTTGAACCTCAGTGAATTTCTTGGAGTCTCTTGGTTTACTTTCACTACGCTGAGAGACATGG	897						
Qy	964	TGTTGAAGTATAGGAGTCAAGAGGTTTGAATCCACCATCACTCAAAATCTGCT	1023						
Db	898	TGTTCAAGTGATAGGGGACCTGAAGAAAGGGTTGAATCCACTCTCGGTTCTGATCTCAT	957						
Qy	1024	ATTTGTGTCCCTTACATGACTACAGCTGTCAAAAATGGCAATTTGCTGGCATCATATC	1083						
Db	958	CTTTACTTCCCTTATATGTCACAGCTGTCAAAAATGGCCCTCATCACTGGCATCATTCG	1017						

RESULT 4  
US-10-437-963-38184  
; Sequence 38184, Application US/10437963  
; Publication NO. US2004012343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua

Qy	1084	ACTTGGCGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATATAAATTAACAATATTGA	1143
Db	1018	TCTCGC-GAAGAGTAGCAGTGGGGAGAGTTTTTGGAGTTTCAAGAACTAACAATAGA	1076
Qy	1144	TGGCAACAAAGAGATGATAGCTATTGGGACCATGAACGCTAGTTGGTTCTTTCACTCTTTG	1203
Db	1077	CGGGNACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGGTTCTTTCACTCTTTG	1136
Qy	1204	CTACTCTCAACAGAGACCATTTTTCGGCTTGGCTGTGAACTATTAACGCTGATGAAGAC	1263
Db	1137	TTACTCTCAACCGGACCATTTTCAAGGTCGCGAGTGAACCTACAAACGCGGTTTGAAGAC	1196
Qy	1264	AGCAGCTTCCAACTAATGTCATTTGCAATGTTGACATTTGTTATTCTTCTGACACC	1323
Db	1197	CGCAATGTCCAACATAGTGATGGCGATTTCGGGTTTANGTTTCACTCTCTTCTTCTCACC	1256
Qy	1324	CTTGTTCATTTACACTCCCTCGTGGTGTCTATCAGCTATTATTCGTATCTGCAATGCTTGG	1383
Db	1257	GCTTTTTCATCTACACACCATCGTCTCTCTCTGCCATCATATATCCGCAATGCTCGG	1316
Qy	1384	ACTCATGATTTAAGAGCAGCCATCATCTATTTAAGTTGACAAAATTTTGACTTTTGGT	1443
Db	1317	ACTCATGCTATCAAGCTGCGCATCTCTCTGAAAGTTGACAAAGTTTCACTCTCTCTCTCGT	1376
Qy	1444	GTGCATGAGTCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCTAGC	1503
Db	1377	CTGCATGAGCCCTACGTTTGGGTCGTATTCGGCAGTTAGAGATTGGAATCTGTCGTAGC	1436
Qy	1504	TATTGTAATCTCTGTACTTTCGGGTACTTCTATTTTATGCAAGGCCAAGGACAATTCGTTT	1563
Db	1437	GGTGGGATATCTATAGCGAGTTTGTCTGTCTTGTCTGGAGGCCAAAACCTGCGGTGAA	1496
Qy	1564	GGGCAACATTCGATATACCGAAATGTTGAGCACTATCAAAATGCAAAACA	1623
Db	1497	GGGAAACATCAACAAACAGCATGATCTATAGAAACACTGAGCAGTACCCATCATCAAGAAC	1556
Qy	1624	TGTTCTCGAATGCTAATTTCTAGAGATTGATGCACTTCTTGTGCAATTTGCAATGCCAGTA	1683
Db	1557	CGTTCCTGGTATCTCATCTTGGAGATTGATGCTCCCACTCTATTTGCTAATGCCAGTTA	1616
Qy	1684	TTTAAGAAAGGATCACAAGTGGATTGATGAAGAAGAAAGAAATTAAGCTTACAGG	1743
Db	1617	CTTTCGTGAAAGAAATCATAGGTGGATTGATGAAGAGGAGAGAGTTAAACAATCAGG	1676
Qy	1744	GGAGCTAGTTTGCAGTATGTTAATTAATGATATGATGCTGTGCAACATTTGATACAG	1803
Db	1677	AGAGAGCAGCTTACAATATATATATCTCGATACTGTCAGTCTGTTGGTAAATATCGACACAG	1736
Qy	1804	TGGAAATAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAGAGAGCTACAGCTTGT	1863
Db	1737	CGGTATTAGCATGATGGTGGAAATTAAGAAAGTCAATTGACAGGAGGCGTTAAAGTTGGT	1796
Qy	1864	TTTGGTCAATCTCTTAAGTGAAGTGAAGAACTGAAGAAATCGAAAGT---TCCAAA	1920
Db	1797	ATTGTCAATCCAAAAGGAGAGGCTGCTGAAGAAATTAACAGATCCAAATTCATCGGTGA	1856
Qy	1921	TCATTTAGGAGAAATGGAATCTATCTGATCTGTTGGAAGGCGCTTGGAGCATGCACTT	1980
Db	1857	TCATTTGGCAAGAGTGGATGTTCTTAAACGTTAGGAGAGCAGTGGAGGCTTGTAGCTA	1916
Qy	1981	CAATCTAGCTGCAAGCAAAACGAAACC	2007
Db	1917	CATGCTTACAGCTTTAAACCGAAACC	1943

APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 38184  
LENGTH: 2372  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_41845C.1  
US-10-437-963-38184

Query Match 34.5%; Score 846; DB 19; Length 2372;  
Best Local Similarity 66.4%; Pred. No. 1.2e-216;  
Matches 1231; Conservative 0; Mismatches 620; Indels 3; Gaps 1;

QY 131 GTCAGGTTCCACGCCACAGCGTCTTCAAGTCTCTAAAGTACTCTTTGAAGGAGACT 190  
DB 152 GTCCGATGCCGCGGCGAAGCGTCTCTGGAGACGCTGGGGGGAACATGAAGGAGACA 211  
QY 191 TTCTTCCCTGATGACCTTTGAGCGAGTCAAGAACAGCCAGCTTCCAAAGAGTTC--- 247  
DB 212 TTCTGCGGACGACCGGTTCAAGGTGTGTCGCGGAGCGCGGTGCGGGCGCGCGG 271  
QY 248 ATGCTTGGCTTCAGTCTCTTCCCATTTTCGAAATGGGCTCCCAAATACACTTTTCAG 307  
DB 272 GCGGCGCGCTCGGTACGTGTTCCCGTTCATGAGTGGGCGCGTGTACACCTTCGGC 331  
QY 308 TTCTTGAAAGCTGACCTCATAGTGGCATCACCATCGCTAGCTTGGCCATTCCTCAGGGC 367  
DB 332 ACCCTCAAGTCCGACCTCATCGCGGCATCACCATTTGCCAGCTCGCCATCCCGCAGGGC 391  
QY 368 ATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTGGACTATATTCGAGCTTTATA 427  
DB 392 ATCAGTACGCCAAGCTCGCAACCTCCCTCCCGTCTTGGGCTCTATTCGAGCTTCGTG 451  
QY 428 CCACCATTTGATTTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACTGTGGCG 487  
DB 452 CCGCGCTGTGTACCGGATGATGGGAGCTCGAGGAGCTCGCGGTGGGAGCGTGGCG 511  
QY 488 GTTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCGGTGATCCCAATGAAGACCCA 547  
DB 512 GTGGCGTGGCTGTGATCGGTCGATGCTGAGCGAGGAGGTGTGCGGCGGAGGACCGG 571  
QY 548 AAGCTTTACCTCACCTGGCTTTACAGCTACATTAATTTGCTGGTGTGTTTTCAGGCTGCC 607  
DB 572 GCGCTGTACCTGACGCTCGCGCTCACCGCACCTTCTTCGCGCGGTGTTTCCAGGCGCTG 631  
QY 608 TTGGGCTGTGTTAGGTTGGGTTGATCGTGGATTTTCTGTACATGCAACCAATAAGG 667  
DB 632 CTGGGGTCTTGAGGCTGGGTTTCATCGTGGACTTCTGTGCGACCGCCACCATCGTCGGG 691  
QY 668 TTCATGGGAGGAGCAGCAGCGTGGTGTGCTGCAGCACTAAATTCGATTTCTTGGCCTT 727  
DB 692 TTCATGGGCGCGCGCCACCGTGGTGTGCTTCAGCAGCTCAAGGCGATGTTTCGSCCTC 751  
QY 728 GAGCATTTACCATGAGCTGATATCATATCAGTATGAGTGGCTCTGTTTTCACCCAAACT 787  
DB 752 GACCACTTCAACCGCCACCGACCTCGTCTCGGTCAATGAGCTCCGCTCTCTCCCAAAC 811  
QY 788 CATGAGTGGAGTGGGAAGTGTGTTAGGATGTCTTCATTTTCTTCTCTCTTCTAGC 847  
DB 812 CACCTCTGGCGATGGGAGAGCGTGTGATGGGTGCGGCTTCTCTTCTCTCTCTCTCATC 871  
QY 848 ACAAGATATTTACAGCAAAAACGACCAAGGTTTTTTTGGTGTGAGCAATGGCGCCATTG 907

DB 872 ACCCGCTTCTCAGCAAGAGGAGGCCAAGGTTCTTCTGGGTATCTGCGCTGCGCCATTG 931  
QY 908 AGTCCGTTATATGGGAAGTCTCTTTGGTTTATTTTCACTCAAGCCGAGAGACGCGTGT 967  
DB 932 GGTCTGTCTATCATCGGAGCCCTCTCTGGTGTACCTCACATGCTGAAAACCATGCAATT 991  
QY 968 GAAGTGATAGGAACTGAAGAAGGGTTTGAATCCACCATCACTCAAAATCTGGTATTT 1027  
DB 992 CAAAGTATGGTTTACCTGAAGAAGGCGCTGAAACCCACCGCTGSCCAAGACCTGAACCTC 1051  
QY 1028 GTGTGCGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087  
DB 1052 TCGTCCCGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1111  
QY 1088 GCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGATATAAAATTTACAATATTTGATGGC 1147  
DB 1112 GCTGAGGGAATGCGGTAGGAAGAGCTTTGCAATGATTTCAAGATTTACATATCGACGGA 1171  
QY 1148 AACAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207  
DB 1172 AACAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231  
QY 1208 CTCACAAAGAGACCAATTTTCGCGTTCGCGTGTGAACATAAAGCTGATGCAAGACAGCA 1267  
DB 1232 CTCACCAAGCGTCCGTTCTCGAGTCCGCGCTCAACTACACAGCCGCGTCAAGACGCG 1291  
QY 1268 GCTTCCAACTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1327  
DB 1292 ATGTGGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1351  
QY 1328 TTCCATTACACTCCCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1387  
DB 1352 TTCCACTACAGCGCTGT 1411  
QY 1388 ATAGATATGAAGAGCCATCCATCTATTTAAAGTTGAACAATTTGACTTTGTTGGTGTGC 1447  
DB 1412 ATCGACTACCGCGCGCGCTCCACCTGTGGCAGGTGGAACAAGTGGACTTCTGCGTCTGC 1471  
QY 1448 ATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCAATGAT 1507  
DB 1472 CTCGGCGCTTACCTCGCGCTGTCTTTCGGCAGGTGAGATCGGCTCTGCTGCGCGCTC 1531  
QY 1508 GTAAATATCTGTACTTCTGGGTACTTCTATTTATTTGAAGGCCAAGGACATTTCTGTTTGGG 1567  
DB 1532 GGGATCTCCATCTCCGCGT 1591  
QY 1568 AACATTTCAAATTTCTGTATATACGAATGTTGAGCACTATCAAAATGCAAAACATGTT 1627  
DB 1592 AACATCCCCAACACGATGATCTACCGCGCGGATGGACCACTACACCGCGCGCAGAGGGTC 1651  
QY 1628 CCTGGAATGCTAAATTTCTAGAGATTTGATGACCAATTTACTTTGCCAATGCCAGCTATTTA 1687  
DB 1652 CCGCGGTGCTCGTGTCTCGCGTGTGATCTCGCCCATCTACTTCAACCCAGCTACCTG 1711  
QY 1688 AGAGAAAGGATCACAAAGTGGATTTGATGAAGAAGAAAGAAATTTAAAGCTACAGGGGAG 1747  
DB 1712 CGTGAGAGGATCGCGGTGGATCGACGACGAGGAGGACAGTGCAGAGGAGAGGCGGAG 1771  
QY 1748 ACTAGTTTCAGATGTTTAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1807  
DB 1772 ATGGGATTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1831  
QY 1808 ATAGTATGCTTTGAAGGTTGAAGATTTACAGAGAGAGAGCTACAGCTTGTGTTTGG 1867  
DB 1832 ACGAGCATGCTGGATGAATCAGGAAGACCTTGGACAGAGGGGCTTTCAGATTTGTGTTG 1891  
QY 1868 GTCAATCTCTGTAAAGTGAAGTGAAGAACTGAACAAATCGAAGTTTCCAAATCATTTTA 1927  
DB 1892 GCGAACCCGGGAGCGAGATCATGAAGATTTGGACAGCTCCAGGTGCTTGGAGCGATC 1951  
QY 1928 GGGAGAAATGATCTATCTGATCTGTTGAAGAGCGGTTGGAGCATGCAACTTC 1981  
DB 1952 GGCCATGATGATCTTCCCAACCGTGGCGGCGGTGGCGAGTGGCGACTTC 2005



```

RESULT 5
US-10-424-599-131117
; Sequence 131117, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Roes Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 39-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131117
; LENGTH: 1518
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89405C.1
US-10-424-599-131117

```

Query Match	32.3%	Score 790.2	DB 18	Length 1518
Best Local Similarity	79.0%	Pred. No. 1e-201		
Matches 1033	Conservative 0	Mismatches 178	Indels 96	Gaps 4
Qy	746	GCTGATATCATATCAGTGATGCGCTCTGTTTTCACCCAAAACATCATGAGTGAGGTGGCAA	805	
Db	24	GCTAAAAGTATTACTTTGAGAGACATTTGTTTTCCCAACGCATGAGTGGAGGTGGAA	83	
Qy	806	AGTGCTGTGTAGATGTGTTTCATTTTCTTCTCTTACGACAAGATTACTTCAGCAAA	865	
Db	84	AGCGCTGTGTGGGTGTGCTTTATTTTCTTCTCTCGTCAAAAGATTACTTCAGCAAA	143	
Qy	866	AAACGACCAAGGTTTTTTTGGGTGTGAGCAATGCGCCATTTGAGTCCGTTTATATTCGGA	925	
Db	144	CGACAGCAAAATGTTCTTCTGGGTGTGAGCAATGSCACCATGAGTCCGTTTATATTCGGA	203	
Qy	926	AGTCTCTTGGTTTATTTTCACTCACGCCGAGAACACGGTGTGAAGTGTAGGAGAACTG	985	
Db	204	AGTCTGCTGTCTATGTCAACACGCTGAGAAACATGGAGTTCAAGTGTAGGAAATTTG	263	
Qy	986	ARGAAGGGTTGATCCACCATCACTCACAAATCTCGTATTGTGTGCGCCTTACATGACT	1045	
Db	264	ARGAAGGGGTGAATCCACCTTCGGCAACGGAATTTGGTGTGTGTGTCACCTTACATGGCC	323	
Qy	1046	ACAGCTGTCAAAACTGGCATTTGCTTGGCATCATATCACTTTCGGGAAGGAATAGCAGTA	1105	
Db	324	ACTGCCATAAAAACCTGGATTCTGCTCACTGGCATATAGCATTTCGGGAAGGAATAGCAGTG	383	
Qy	1106	GGAAGAAGCTTTGCAATGTATAAAANTACAATTTGATGGCAACAAGAGATGATAGCT	1165	
Db	384	GGAAGAAGCTTTGCAATGTATAAAANTATCATATTGATGGCAACAAGAGATGATAGCT	443	
Qy	1166	ATTGGGACCATGAACAGTAGTTGGTTCTTTTTCACCTCTTGTACCTCAACAAGGACCACTTT	1225	
Db	444	ATTGGAACCATGAATATTTTCGGATCGTTCACTTCTTGTCTATCTCACCATGGACCACTTT	503	
Qy	1226	TCGCGTTTCGGGTGGAACATAAAGCTTGGATGCAAGACAGCAGGTTTCCAAATATTAATG	1285	
Db	504	TCGCGGTTCAGCAGTGAACATAAATGCTGGGTGCAAGACAGCTGCATCCAAACATCATTAATG	563	
Qy	1286	TCACCTTCGAGTAATGTTGACATTTGTTATTTCTCTGACACCTTTGTTTCCATTACACTCCCTG	1345	
Db	564	GCAATTCGAGTAATGTTGACATTTGTTATTTCTTAACACATTTGTTTCCATTACACTCCCTG	623	
Qy	1346	GTGTGCTATCAGCTATTATTCGTAATCTGCAATGCTTGGACTCATAGATTATGAAGACGCC	1405	
Db	624	GTGTGCTATCAGCATTATTTGTTATCTCGAATGCTTGGCCTCATAGATTATGAAGACGCC	683	

Qy	1406	ATCCATCTATTAAAGTTTGACAAATTTGACTTGTGTGTCATGAGTGCATACATTGGC	1465
Db	684	ATCCATCTATTTAAAGGTCGACAAATTTGACTTTTGTGTGTCATGAGTGCATGCTTTGGC	743
Qy	1466	GTGGTCTTTTGGCAGTGTGTGAAATTTGGCGTTAGTCATAGCTATTGTAATCTGTACTTCTGG	1525
Db	744	GTGGTCTTTTGGCAGTGTGTGAAATTTGGCGTTAGTCATAGCTATTGCGATATCTGTACTTCTGG	803
Qy	1526	GTACTTCTATTTATTGCAAGCCCAAGACACATTCGTTTTTGGGCAACATTTCCAAATTTCTGTG	1585
Db	804	GTACTTCTATTTATTGCAAGGCCA--GACATTTGTTTTGGGAACATTTCCAAATTTCTGTG	861
Qy	1586	ATATACCGAAATGTGTAGCACTATCAAATGCAAAAACATGTTCTCTGGAATGCTTAATCTTA	1645
Db	862	ATATACAGAAATGTGTAGCACTATCCAAATGCAAAAACATGTTCTCTGGAATGCTTAATCTTA	921
Qy	1646	GAGATTGATGACCAATTTACTTTGCCAATGCCAGCTATTAAAGAGAAAGGATCACAGG	1705
Db	922	GAGATTGATGACCAATTTACTTTGCCAATGCAAAAGATCAAGC-----TGCT	979
Qy	1706	TGATTTCATGAAGAAGAAAGAAATTTAAAGCTACAGGGGAGACTAGTTTCAGCTATGTT	1765
Db	980	TGATTTCATGAAGAAGAAAGAAATTTAAAGCTACAGGGGAGATCAAGC-----TGCT	1015
Qy	1766	ATAATTGATATGAGTGTCTGTGGAAACATTTGATACAAAGTGAATAGTATGCTTGAAGAG	1825
Db	1016	ATGTTTCATATGAGTGTCTGCTGTA-----	1041
Qy	1826	GTGAGAAGATTTACAGAGAGAGAGACTACAGCTTGTGTTGGTCAATCTGTAAGTGAA	1885
Db	1042	-----TGCTTGGGGAGATTTCCAAGAAAGCGAG	1067
Qy	1886	GTGATGAAGAAACCTGAACAAATCGAAGTTTCCAAATCAATTTAGGGAAGAAATGGATCTAT	1945
Db	1068	GTGATGAAGAAAGCTGAACAAATCCAGTTTCTTGATGAACTGGGGCAAAATGGATCTAT	1127
Qy	1946	CTGACTGTTGAAGAGCGCTTTGGAGCATGCAACTTCAATCTTACGTGCAAGCAAAACGAAAC	2005
Db	1128	CTGACAGTTGAAGAGCGCTTTGGTGTGCTGCAACTTCATGCTCCATTCCTACAAACCGAAAC	1187
Qy	2006	CCAAAGAAAGATGAAACAGAGGGTTTGAACAAATGTGTGACTGATCTCA	2052
Db	1188	CCCATGAAGATGAATCAGAAGGTTTGAACAAATGCCCTAAATAAAATAA	1234
RESULT 6			
US-10-425-115-50601			
; Sequence 50601, Application US/10425115			
; Publication No. US20040214272A1			
; GENERAL INFORMATION:			
; APPLICANT: La Rosa, Thomas J.			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants			
; FILE REFERENCE: 38-21(53222)B			
; CURRENT APPLICATION NUMBER: US/10/425,115			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 369326			
; SEQ ID NO 50601			
; LENGTH: 2727			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; OTHER INFORMATION: Clone ID: MRT4577_146143C.1			
US-10-425-115-50601			
Query Match 27.7%; Score 677.4; DB 20; Length 2727;			
Best Local Similarity 60.8%; Pred. 3.6e-171; Indels 3; Gaps 1;			
Matches 1123; Conservative 0; Mismatches 721;			
Qy	122	GTGCACCAAGTCGAGGTTTCACCCGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTG	181

Db 300 GGGTACAAAGTCCGGCTCCACCGGAGAAACCTCTCTCCCGAGATCTCTGACGGGTG 359  
Qy 182 AAGGAGACTTTCTTCCCTGATGATCCCTTTGAGCGAGTTCAGAACAGCAGCTTCCAG 241  
Db 360 AAGGAGACGTTCTTCGCGGACGACCCGCTCGGCGAGTACAGGACACCGACCGTCCAG 419  
Qy 242 AAGTTCAATGTTGGCCCTTCAGTTCTTCTTCCCAATTTTCCAAATGGGCTCCCAATACACC 301  
Db 420 AAGATCTGGCTCGGCTCGAGACATCTTCCGGTCTCGAATGGAGCGCATTTACTCC 479  
Qy 302 TTTCAAGTTCTGAAAGTCACTCATAGCTGGCAATCCATCGTACGTAGTTGGCCATTCCT 361  
Db 480 CTGGCAAGTTCAAGGGTGAATTCATCGCTGGCTCACCATTCGACGCTCTGCATACCC 539  
Qy 362 CAGGCAATCAGTTATGCAAGCTCGCAACTCTCTCCCAATTTTCCAAATGGGCTCCCAATACACC 421  
Db 540 CAGGACATCGGTAATTTAAGCTTGTCTTCTTCCAGCAGAGTGGGACTATACAGTAGC 599  
Qy 422 TTTATACCAATGATTTATGCGATGATGGGTAGCTCGAGGATTTGGCAGTGGGACT 481  
Db 600 TTGCTGGCTCTGATATACGAGTATGGGAGCTTCAGGACATAGGCCATCGGTCCA 659  
Qy 482 GTGCGGTTGGATCGCTTCTGATGGGTTGATGTGAGTAATGCCGTTGATCCCAATGAA 541  
Db 660 GTGCGGTTGGTCTGCTGCTGCTCGGTACTCTCTTCCAGAAATGAGATCGAACCCCAAGACA 719  
Qy 542 GACCCAAAGCTTTACCTCACCTGGCTTTCACAGCTACATTTTGGTGGTGTGTTTCAG 601  
Db 720 CACCCGCTCGAGTACAGGCGCTAGCCTTTCACAGCAACCTTTTTCGACGGGTCACCTCAG 779  
Qy 602 GCTGCTTGGGTTGTTTGGGTTGATCGTGGATTTTCTGTCACATGCAACCAATA 661  
Db 780 GCAGCCCTCGATTTCTCAGGCTAGGTTTATCATAGATTTCTTGCTCATGCTGCCAT 839  
Qy 662 ATAGGGTTCAATGGGAGGAGCAGCAAGTGGTGTGCTGCAGCAACTTAAATCGATTCTT 721  
Db 840 GTCGGATTCATGGCGGCTGCCATCACCAATTTGCGCTTCAGCAGCTGAAGATTCCTT 899  
Qy 722 GGCCTTGAGCAATTCACCCATGAGCTGATATCATATCATGATGCGCTCTGTTTTCACC 781  
Db 900 GGAATTTGCAAACTTCACCAAGAAATCCGATATTTGATCTGTTTGAATCAGTTTGGGA 959  
Qy 782 CAACACT---CATGAGTGGAGTGGAAAGTGTGTTAGGATGTGTTCTCATTTTCTTC 838  
Db 960 AATGTTCAACATGGGTGGAATGGCAGCAATCTGATAGGCGACCTTCTCTGGCAATC 1019  
Qy 839 CTCCTTAGCAACAAGATACTTCAGCAAAAACGACCAAGGTTTTTTGGGTGTCAGCAATG 898  
Db 1020 CTTCTGGTTGCCAAGTACATTTGGAAAGGAATAAAGCTCTTCTGGGTGTCGCAATC 1079  
Qy 899 GCGCCATGACGTCGGTTATTTATTTGGGAAGTCTCTGGTTTATTTCTACACGCGGAGAAG 958  
Db 1080 GCACCTCTCACTTCGGTGATCATATCCACATTTTGTGTACATCACTCTGTCAGATAAG 1139  
Qy 959 CAGGTTGGAAGTATAGGAACTGAAGAGGTTTGAATCCACATCACTCACAAT 1018  
Db 1140 CATGGCGTTGCAATTTGCAAGACATAAGAAAGGATCAACCCACTTCAGCTAGTCTA 1199  
Qy 1019 CTGGTATTTGTGTCGCTTACATGACTACAGCTGTCAAACTGGCATTTGCTGGCATC 1078  
Db 1200 ATATACTTACCGGCCCATCTTGGGACAGGATTCAAAATGGGATAGTAGTGGAAATG 1259  
Qy 1079 ATATCATCTTGGGAAGGAATAGAGTAGGGAAGCTTTGCAATGTATAAAAATTAAT 1138  
Db 1260 ATAGGCTTAACGGAAGGATTTGCAATTTGGAAGAACATTTGACGCTTCAAGGATTAACCG 1319  
Qy 1139 ATTGATGGCAACAAAGAGATGATAGTATTTGGACCATGACAGTGGTCTTTTTCACC 1198  
Db 1320 ATAGATGGGAACAAGAAATGATGGGCTTAGGACCATGAAACATTTGTTGTTCACTGACT 1379  
Qy 1199 TCTTGTCTACCTCAACAGGACCATTTTTCGGTTCGGCTGTGAACTATAAACGCTGGATGC 1258

Db 1380 TCTTGTACGTAGCCACAGGTTCTTTCTCGCGTCAGCAGTTAAATTACATGGCTGGCTGC 1439  
Qy 1259 AAGACAGCAGCTTCCAAATTAATAATGCTCACTTGCAGTAATGTTGACATTTGTTATCCCTG 1318  
Db 1440 AAMACAGCAGTTCAAACGTTGTTAATGTCATTTGTCGTAATGCTTACTTTGCTACTGATC 1499  
Qy 1319 ACACCCCTTGTTCATTTACACTCCCTCGTGGTGTCTATCAGCTATTTATCGTATCTGCAATG 1378  
Db 1500 ACCCATTTGTTCAAGTACACTCCAAATGCATCTTTCTCGATCATCATATCAGCAGTG 1559  
Qy 1379 CTTGGACTCATAGATTTAAGCAGCCATCCATCTATTTAAGGTTGACAAATTTGACTTT 1438  
Db 1560 CTTGGTTAATTTGACTATGAATCGGCTTACTTATCTGAAAGTTGACAACTGGACTTT 1619  
Qy 1439 GTGGTGTGATGAGTGCATACATTTGGCGTGTCTTTGGCAGTGTGCAAAATGGCTTAGTC 1498  
Db 1620 CTAGCATGATGGAGCATTTTGTGGAGTCATATTTTCAATCGGTGAGTATGGCTTGCTC 1679  
Qy 1499 ATAGCTATTTGTAATCTGTACTTTCGGGTACTTCTATTTTATTTGCAAGGCCAAGCAGATTC 1558  
Db 1680 AITGCGTTGCAATATCTCTTCTTAAATTTCTTCCAAAGTAAACACGCGCAAGACAGTT 1739  
Qy 1559 GTTTTGGCAACATTCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCA 1618  
Db 1740 TTACTTTGAAACCTTCCAGAACTATATATACAGGAATGTAGAACAGTATCTGATGCT 1799  
Qy 1619 AAACATGTTCTCGAATGCTAATTTCTAGAGATGTAGACCAATTTACTTTGCAATGCC 1678  
Db 1800 ACCAAGTTCCAGGGGTGCTGATTTGTAGAGTGGATCTCAGCTATATACTTCAAACTCT 1859  
Qy 1679 AGCTATTTAAGAAAGGATCAAGGTGGATTTGATGAAGAAAGAAAGAAATTAAGCT 1738  
Db 1860 AACTATTTAAGAGAGATCTGAGTGGCTAAGAGATGAGGAGGAGCAACAGGAC 1919  
Qy 1739 ACAGGGAGACTAGTTTGCAGTATGTTAATTAATGATATGAGTGTCTTTGGAAACATGAT 1798  
Db 1920 CAGAAGTTAAACAAAACCTGAGTTTCTAAATTTGTGACCTGTCTCTGTAATGATATCGAC 1979  
Qy 1799 ACAGTGGAAATAGTATGCTTGAAGAGTGAAGAAATTTACAGAGAGAGAGCTACAG 1858  
Db 1980 ACAAGTGGAAATCAATGCTTTGGAGGAGTTGGGAAAGCTCTTGAAAAACGCAAAATTCAG 2039  
Qy 1859 CTTGTTTTCGTCATCTCTAAGTGAAGTGAAGTGAAGAACTGAACAAATCGAAATTTCCAA 1918  
Db 2040 CTGGTTCTTACCAATCCCGGCGGGGTGATCCAGAGCTCCGCTCAGGAAATTCAG 2099  
Qy 1919 AATCAATTTAGGAAGAAATGGAATCTATCTGACTGTTTGAAGAGCGCT 1965  
Db 2100 GACATGTTGGTGAAGACAAACATATTCCTCAGGTCGGCGAGCTGT 2146

## RESULT 7

US-10-437-963-77012  
; Sequence 77012, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Roba, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 77012  
; LENGTH: 4390  
; TYPE: DNA

; ORGANISM: Oryza sativa									
; FEATURE:									
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76952C.1									
US-10-437-963-77012									
Query Match 27.4%; Score 670.4; DB 19; Length 4390;									
Best Local Similarity 60.3%; Pred. No. 3.7e-169;									
Matches 1109; Conservative 0; Mismatches 731; Indels 0; Gaps 0;									
QY	126	ACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTGAAGG	185						
DB	504	ACAATGTCGGCGCGCCGCGAGAGAACCTCTCGCGGAGATTCCGCGGACCGTGAAGG	563						
QY	186	AGACTTTCTTCCCTGATGACCCCTTTGAGGAGATTCAAGAACAGCCAGCTTCCAAAGAGT	245						
DB	564	AGACGTTCTTCTCCGACGAGCCGATGCGGCGGTACAAGGACCGCCAGGTCGAGGAAGC	623						
QY	246	TCAATGCTTGGCCTTCAGTTCCTTCTCCCAATTTTTCGAATGGGCTCCCAAAATACACCTTTC	305						
DB	624	TATGGCTCGCCTTGCAGCACGTCTTCCCGGTGTTTCGAATGGGCGAGACAATACACCTTCG	683						
QY	306	AGTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTTGGCCATTCCTCAGG	365						
DB	684	CCAAGTTCAGGGCGACCTCATTCGCGGCTCACCTTGCAGGCTCGTCATACCTCAGG	743						
QY	366	GCATCAGTTATGCAAGCTCGCAACCTCCCTCCAAATCTTTGGACTATATTCGAGCTTTA	425						
DB	744	ACATCGGCTACGGAGCTTGCTTAACCTGCCACGAGATTGGGCTGCACAGTAGCTTCG	803						
QY	426	TACCACCAATTTATGCGATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGAGCTGTGG	485						
DB	804	TCCCGCATTTGATACGCTCTGATGGGCACCTCAAGGGAGCTAGCAATGGGTCCAGTGG	863						
QY	486	CGGTTGGATCGCTTCTGATCGGTTGCAATGTTGATTAATGCCGTTGATCCCAATGAAGACC	545						
DB	864	CGGTCACTCACTGCTGCTTGTACTCTCTCCAGGAGGAGATTGACTCAAGAGAAAGAAC	923						
QY	546	CAAAGCTTTACCTCCACCTGGCTTTCACAGCTACATTTATTTGCTGGTGTTTTTTCAGGCTG	605						
DB	924	CGTAGATTACAGACGGCTCGCTTTCACAGGACCTTCTTTGCAGGAGTTCACACAGGCG	983						
QY	606	CTTGGGTCTGTTTAGGTTGGGTTGATCGTGGAATTTCTGTCTCATGCAACCAATAATAG	665						
DB	984	CGCTGGGTTTCTGCAGGCTAGGGTTTCATCATAGCGTTCTCTCTCATGCTGCCATCATCG	1043						
QY	666	GGTTTCATGGGAGGAGCAGCCAGCGTGTGCTCTGCAGCACTAAATCGATTCTTCGCCC	725						
DB	1044	GATTCATGGCGCGCGCCATCACCTATGCTCTTTCAGCAGCTTAAAGGCTTCTTTGGAA	1103						
QY	726	TTGAGCATTTTACCCATGGAGCTGATATCATATCAGTATGCGCTCTGTTTTACCCAAA	785						
DB	1104	TTGCAAACTTCAACAGAGACTGACATCATCTCGGTCATGAATCAGTCTGGGGAATG	1163						
QY	786	CTCATGAGTGGAGTGGGAAGTCTGTGTTAGGATGTCTTCATTTTCTTCTCCTTA	845						
DB	1164	TTCAACATGGGCAATGGAACTGGCAGACAATAATTGATCGAGCATCATTTTGGCATTC	1223						
QY	846	GCACAAGATACTTCAAGAAAAACGACCAAGGTTTTTTGGGTGTGAGCAATGCGCCAT	905						
DB	1224	CTCCTGTTGCCAAGGCAAGAGAAACAAGAAGCTCTTCTGGTCCGAGCAATGCAACCAC	1283						
QY	906	TGACGTCGGTTATATTGGGAAGTCTCTTGGTTTATTTTCACTACGCGCGAGAACGCGTG	965						
DB	1284	TCAATTCGGTGTATTCATTTCAACTTTGTTGGTCTACATCACTCGTGTGACAAACAGGCG	1343						
QY	966	TTGAAGTGTAGGAGACTGAAGAAGGTTTGAATCCACCATCACTCACAAATCTGGTAT	1025						
DB	1344	TGCAATTCGTAAAAAAGCTCAAGAAAGGCATCAATCCACTTCAGCTAGCTGATTTT	1403						
QY	1026	TTGTGTGCGCTTACATGACTACAGCTGTCAAAACTGGCATTTGCTGGTCATCATATCAC	1085						
DB	1404	TCACTGCCCACTATTGCTCAAGAGATTCAAATTTGGAGTAGTAGTGGAAATGATAGCC	1463						

QY	1086	TTGCGGAAGGAATAGCAGTAGGAAGAGCTTTTGCATATGTATAAAATTTACAATATTGATG	1145
DB	1464	TTACGGAAGCAATGTCAGTTGGAAGAACATTTTCTGGGATTTGAACGATTTACAGATAGATG	1523
QY	1146	GCAACAAAGAGATGATAGCTATTGCGGACCATGAAAGTGTGGTTCTTTTCACTCTTGCT	1205
DB	1524	GGAACAAAGAAATGTTGCTCTTAGGAACCATGAAATGTGGTCTGGTTCAATGAGCTTGTCT	1583
QY	1206	ACCTCAACAGGAGCCATTTTCGCTTGGCTGTGAACATATAACCTGTGATGACAGACAG	1265
DB	1584	ATATAGCCACAGGTGTTTTCGACGATCAGCAGTCAATTCGATGGCTGGAGTAAAAACAC	1643
QY	1266	CAGCTTCCAAATATATATGTCACCTTGCAGTAAATGTTGACATTTGTTATTCCTGACACCT	1325
DB	1644	CAATGTCCAATATTTGTTATGCACTAGTATTTCTTGGCACTCTGCTGGATCATCTCAT	1703
QY	1326	TGTTCCATTTACACTCCCTGGTGGTGTATCAGCTATTTATCGTATCTGCAATGCTTGGAC	1385
DB	1704	TGTTCAAGTCACGCCCAATGCCACCATTTCTTCCATCATATATCAGCAGTGTGGCC	1763
QY	1386	TCATAGATTTGAAGCAGGCATCCATCTATTTTAAGGTTGACAAATTTGACTTTGTGGTGT	1445
DB	1764	TATTTGACTTTGAAATCAGCCTACCTTATCTGGAAGTTTGAATAGTTGACCTTCATGGCT	1823
QY	1446	GCATGAGTGTCATACATTTGCGGTGGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATAGCTA	1505
DB	1824	GCTTGGGGCATTCCTTGGAGTAATTTTCTGCTGGAGTATGGCTTGTCTCATTTGGCG	1883
QY	1506	TTGTAAATATCTGACTTCGGGTACTTCTTATTTTATTCGAAGGCCAAGGACATTCGTTTTGG	1565
DB	1884	TTGTAAATATCACTAATCAAGTTCTGCTCCATGTAAACGCGCAAGGACAGCTTTACTTTG	1943
QY	1566	GCAACATTTCCAAATTCGTGTATATACCGAAATTTGTGAGCACTATCAAAATGCAAAACATG	1625
DB	1944	GCAACCTTCCAAAGACGATTTATCTATAGAAATTTGTAACAAATATCCAGAAGCTACCAAGG	2003
QY	1626	TTCTCTGGAATGCTAATCTTAGAGATTGATGACCAATTTTACTTTGCTCAATGCCAGTATT	1685
DB	2004	TGCCAGGATGCTAATTTGTAGAGTGGACTCAGCAATATCTTCAAAACTCCAATTATG	2063
QY	1686	TAAGAGAAGGATCAAGGTGGATTGATGAAGAAGAAAGAAATTTAAAGCTACAGGGG	1745
DB	2064	TTAAAGAAGAATGCTGAGATGGCTGAGAGATGAGGAAGAACATCAAAAGGAACAGAAGT	2123
QY	1746	AGACTAGTTTGCAGTATGTTATTAATGATGAGTGTGCTGTTGGAAACATTTGATACAGTG	1805
DB	2124	TACCAAAATTTGAGTTTCTGATTTGTGACCTATCTCTCTGTAATGATATTGACACAAGTG	2183
QY	1806	GAATAAGTATGCTTCAAGAGGTGAAGAAGATTACAGAGAGAGAGAGCTACAGCTTGTTT	1865
DB	2184	GAATCCATGCTTCAAGAGTTTGTGAGGACACTTTGAAAGCGCCAGATTGAGCTGATTT	2243
QY	1866	TGCTCAATCTCTTAAGTGAAGTGAAGAAGAACTGAACAAATTCGAAGTTTCCAAATTCATT	1925
DB	2244	TCGCCAATCTCTGGGCGGCTGTGATCCAAAGCTCCGGTCAGCAAAATTTACAGAGCTCA	2303
QY	1926	TAGGGAAGAAATGATCTATCTGACTGTTTGAAGAGCCGT	1965
DB	2304	TTGGTGAAGAAAGATATATGCTGACAGTTGGTGGACGCGGT	2343

RESULT 8

US-10-424-599-65103  
; Sequence 65103, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21 (53223)B

! CURRENT APPLICATION NUMBER: US/10/424,599  
! CURRENT FILING DATE: 2003-04-28  
! NUMBER OF SEQ ID NOS: 285684  
! SEQ ID NO 65103  
! LENGTH: 2366  
! TYPE: DNA  
! ORGANISM: Glycine max  
! FEATURE:  
! OTHER INFORMATION: Clone ID: PAT\_MRT3847\_29800C.1  
US-10-424-599-65103

Query Match 26.9%; Score 658.6; DB 18; Length 2366;

Best Local Similarity 60.5%; Pred. No. 3.9e-166;  
Matches 1120; Conservative 0; Mismatches 724; Indels 7; Gaps 2;

QY	77	AGTGTAGATATGAGTACCTTTGGGCATCAACAACTTGGAGAGATGCGACCAAGTCGAG	136
DB	132	AATGCAATGATATGATCAATCTCATTCATGATAGATGTCCATGGAAGTCACCAAGTTGTG	191
QY	137	GTTCACCGCCACAGCGGTTTTTCAAGTCTCTAAAGTACTCTTTTGAAGGAGACTTTCTTTC	196
DB	192	CCACCACTTCAAGAGCACCCTTCAGAACTCAAGGTAGGCTCAAGGAATCTTCTTC	251
QY	197	CTGTATGACCTTTTGAAGCAGTTCAAGAAACAAGCCAGCTTCCAAAGTTCATGCTTGC	256
DB	252	CCTGATGATCTCTCGCCCAATTCAGGGACAACTCTTAAAGAGAAATCGATCTTGA	311
QY	257	CTTCAGTCTCTTCCCAATTTTGAATGGCTCCCAATATACACCTTCAGTCTTCTGAAA	316
DB	312	GCTCAATATGTGTCCCTATTTCTCAATGGGGTCTTAAGTATAATCTCAACTCTTCAAA	371
QY	317	GCTGACCTCATAGCTGGCATCACATCGCTAGCTTGGCCATTTCTCAGGSCATCAGTTAT	376
DB	372	TTTGACCTTGTCTGGCCCTCACTATGTGATGCTTGGACATCCCGCAGGGAATGATTAI	431
QY	377	GCCAACTGCCCAACCTCCCTCAATTTCTTGAAGTATATTCAGAGCTTTATACACCAATG	436
DB	432	GCTTAGCTTGCAGTCTTCTCCAAATTTAGGACTTTATCTAGTTTGTCCACCACTT	491
QY	437	ATTATGCGATGATGGTGTAGCTCGAGGATTTGGCAGTGGGAGCTGTGGCGGTGATCG	496
DB	492	GTCTATGCTGTTCTTGGAACTCAAGGACCTTGCAGTTGGACCTGTTCTATTTGCTTCT	551
QY	497	CTTCTGATGGTTCGATGTGATGATGCGTTGATCCCAATGAAGACCCAAAGCTTTTAC	556
DB	552	CTTGTATGGATGCCATGTGATCAGGAAGTGTCTCCCAACACAGATCCAAATTCGTTT	611
QY	557	CTCCACTGGCTTTCAAGCTACATTAATTTGCTGGTGTGTTTTCAGGCTGCTTGGGCTCG	616
DB	612	CTTCAGCTAGCTTTCACTTCAATTAATTTGCGGTCTCTTTCAGCTTTGCTTGGAAATC	671
QY	617	TTTATGTTGGGTTGATTCGTGGATTTCTGTCAATGCAACCAATATAGGGTTCAATGGA	676
DB	672	CTAAGGCTAGGCTTCATAATTTGATTTCTATCTAAGGCCATCTTATTTGGGTTCTGGCT	731
QY	677	GGAGCAGCCAGGTGTGTCTGAGCAACTAAATCGATTTCTGGCCCTTGAAGCAATTC	736
DB	732	GGAGCTGCTATTAATTTGCTCACTGCAACAGCTCAAGAGCTGCTTGGAAATCAACATTC	791
QY	737	ACCCATGGAGCTGATATCATATCAGTGTGCTGCTGTTTTCACCCAACTCATGAGTGG	796
DB	792	ACTAATTCAGATGGGCTGATCTCTGTATTAAGTCTTCTGTTTTCACATATATCATGATGG	851
QY	797	AGGTGGAAAGTGTCTGTGTAGATGTCTTCAATTTCTTCTCTCTTGAACAAGATAC	856
DB	852	TCATGGCAACAATATTTGATGGGATTTGCTTCTTGTGTACTACTACTATTAGCAAGACAC	911
QY	857	TTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTAGCAATGGCGCCATGTAGCTCCGTT	916
DB	912	GTTAGCATAGGAAACCAAACTATTTCTGGGTCTCAGCTGGAGCTCTCTCTTATGTGTGTC	971
QY	917	ATATTGGGAAGTCTCTTGTGTTTATTTCACTCAAGCCGAGACGACGCTGTGAAGTATA	976

## RESULT 9

US-10-425-115-182207  
; Sequence 182207, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.

DB	972	ATCATCTCTACCTCTCTGGTTTTTGCATAATTAAGGCTCAAAATCATGGCATCAGTCGAT	1031
QY	977	GGAGAACTGAAGAAGGTTTGAATCCACCATCACTACAAAATCTGTATTTGTGTCGCT	1036
DB	1032	GGAAATTTGCAACAAGGAATAAATCTCATCATGGAATATGTGCTCTTTTCATGGAAT	1091
QY	1037	TACATGACTACAGCTGTCAAACTGGCATTTGCTTGGCATCATATCACTTGC-GGAAGG	1095
DB	1092	CACCTAGGCTAGTTTATGAAAACAGGCTTATCACCGGATTTTGTCCCTAAACAGGAAG	1151
QY	1096	AATAGCAGTAGGAAGAAGCTTTGCAATGTATATAAATTAACAATATGATGGCAACAAAG	1155
DB	1152	TATTCAGTAGGAAGGACATTTTGCAGCTCTCAAAACTCAAAAGTGAATGGAATAAGGA	1211
QY	1156	GATGATAGCTATTTGGGACCATGAACGTAGTTGGTTTTCTTCACTCTTGTCTCACTCAAC	1215
DB	1212	AATGATGGCAATGGGTTTATGAATTTGTTGGCTCTCTTCACTCTCTGCTATGTTACAAC	1271
QY	1216	AGGACCATTTTTCGGCTTGGCTGTGAACCTATAACGCTGGATGCAAGACAGAGCTTCCAA	1275
DB	1272	AGGTGCTTCTCTCGTCAAGCTTAAACAATCAAGCGCAAAACAGCTGTGTCAAA	1331
QY	1276	CATTATATGCTCACTTGCAGTAAATGTTGACATTTGTTATTCCTGACACCTTGTTCATTA	1335
DB	1332	TGTAGTGTGTCTGTGACAGTCAATGTCACCTCTTTTCTCATGCCATTTGTTCAATA	1391
QY	1336	CACTCCCTGGTGGTCTATCAGCTATTAATCGTATCTGCAATGCTTGGACTCATAGTTA	1395
DB	1392	CAGGCTAATGCTGGTGGGCGCAATCATAGTCACAGCAGTAATTTGGCTCATCGATCT	1451
QY	1396	TGAAGCAGCCATCACTATTTAAGGTTGACAAAATTTGACTTTTGGTGTGATGAGTGC	1455
DB	1452	CCCCGCTGCTGTAACTTTGGAAGATCGACAAATTCGATTTTGTGATGATGACTGC	1511
QY	1456	ATACATTTGGCGGTCTTTTGGCAGTGTGAAATTTGGCTTTAGTCACTAGTATTTGTAATC	1515
DB	1512	TTTCTTAGTGTCTTTTCTCATCTCTGTCCAAGGAGCTTGTCTTGTGTTGGATTATC	1571
QY	1516	TGTACTTGGGTACTTCTTATTTATTCAGGCGCAAGGACATTCGTTTGGGCAACATTC	1575
DB	1572	AACTTTGAAGATACCTTTGCAAAATTCAGAGACCGAAACAGTGTGTTGGGAAGATACC	1631
QY	1576	AAATTTCTGTATATACCGAAATGTTGAGCACTATCAAAATGCAAAACATGTTCTTGGAAAT	1635
DB	1632	AGGAACAGACATATATAGAAATCTTGATCAATACAGGAGCTGTGAGATACCTGGAT	1691
QY	1636	GCTAATTTAGAGATTTGATGCAACCAATTTAATTTGCCAATGCCAGCTATTTAAGAGAAAG	1695
DB	1692	TCTTATTTTAAAGCAATGAGGCTCCCATCAATTTTGCACATCAATATCTCAACGAGAG	1751
QY	1696	GATCAAGGTGGATTTGATGAAGAAGAAGAATTAAGCTACAGGGGAGACTAGTTT	1755
DB	1752	AACTTAAAGATGGATTTGAAGAAGAAGAACAATAA-----GGAACAATTAAGCT	1805
QY	1756	CGAGTATGTTAATTTGATATGAGTGTGTTGGAACAATTTGATACAAGTGAATTAAGTAT	1815
DB	1806	TCGATTTCTTAGTATTTGGAATTTGAGCTGTGAGTGTGTTGACAGAGTGAATCTCACT	1865
QY	1816	GCTTGAAGAGGTGAAGAAGATTTACAGAGAGAAGAGCTACAGCTTGTGTTGGTCAATCC	1875
DB	1866	TTTCAAGGATTTGAAGCAACACTCGAAAAGAAAGGTGTGAGCTTGTGTTGGTCAATCC	1925
QY	1876	TGTAAGTGAAGTGAAGAAGAACTGAAACAAATCGAAGTTTCCAAATCATTT	1926
DB	1926	TCTTGTGAGGTCTATAGAAAAGCTTTAAAAAAGCAGATGAAGTAAATGATTT	1976



## RESULT 10

US-10-437-963-77005  
; Sequence 77005, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 77005  
; LENGTH: 2780  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRP4530\_76946C.1  
US-10-437-963-77005

Query Match 26.4%; Score 647.2; DB 19; Length 2780;

Best Local Similarity 59.8%; Pred. No. 5.1e-163;

Matches 1103; Conservative 0; Mismatches 738; Indels 3; Gaps 1;

QY	125	CACCAAGTCGAGGTTCCACGCGCACAGCCGTTTTCAGAGTCTCTAAAGTACTCTTTGAAG	184
DB	457	CACAAGTCGAGTTCCACGCGAGAGAGCTCATCGAGGTTCAACGCGCGTGAG	516
QY	185	GAGACTTTTTCCTGATGACCTTTGAGGCGAGTTCAGAAACAGCCAGCTTCCAAAGAG	244
DB	517	GAGACGTTCTCGCCGACGACCGTTGCGGCGAGTCAAGGACGCGGATGTCAAAGAG	576
QY	245	TTACGCTTGGCTTCAGTCTTCTTCCCAATTTTCCGATGGCTCCCAATACACCTTT	304
DB	577	GTGTGATCAGCTGCAGAACTTCTTCCGCGTGTGAGTGGGCGCGCACTACACCTTC	636
QY	305	CAGTTCTTCAAGAGCTCAGCTCAGTGGCATCAGCTCAGCTGAGCTTGGCCATCTCTCAG	364
DB	637	CGCAAGTTCAAGGCGGACCTGCTCTCCGCGCTCACCATTTGCCAGCTCTGCAATACCCAG	696
QY	365	GGCATCAGTTATGCCAAGCTCGCCACCTCCCTCCCAATTTCTTGGACTATATTCGAGCTTT	424
DB	697	GACATCGGTTATCGGAAGCTTGCTGGCCCTATTACCAATTTATGGACTGTACAGCAGCTTC	756
QY	425	ATACCAACCAATTGATTTATCGGATGATGGTAGCTCAGGAGATTTGGCAGTGGGACTGTG	484
DB	757	GTACCGCTTTGATATACGCGATGATGGGAAGTTCCAGGGATATAGCCATTGGTCCAGTG	816
QY	485	GGGTTGGATCGCTTCTGATGGTTGCGATGTTGAGTAATGCGGTGATCCCAATCAAGAC	544
DB	817	GGCTGTTTGGTGTGCTTGGCACTCTCTCCAGATGAGTTGATCCCAAGAGAAAT	876
QY	545	CCAAAGCTTTTACCTCCACCTGGCTTTTCAGAGTACATTAATTTGCTGGTGTCTTTCAGGCT	604
DB	877	CAGGAAGATACACCGGCTGGCTTCACTGCAACGTTCTGCTGGGTCAACCCAGCA	936
QY	605	GGCTGGGCTGTTTATGGTGGGTTGATCGTGGATTTTCTGTCATGCAACCAATAATA	664
DB	937	GTGCTCGAATTCCTCAGGCTAGGTTTATCATAGAGTTCTTGTCCCATCTGCTCCCATCGTT	996
QY	665	GGGTTTCAAGGAGGACCGGCTGGTGTCTGCGCAACTAAATCGATTTCTGCG	724
DB	997	GGATTCATGGGAGGCGGCCATACCAATTTGCCCTTCAGCAGCTTAAAGGCTTCTTGGGA	1056
QY	725	CTTGAGCATTTCAACCATGAGCTGATATCATATCAGTGATGCGCTCTGTTTTCACCCAA	784

DB	1057	ATTGCAAACTTCCAAAGAAGACTGACATCATCTCGGTATGAAATCAGTCTGGGGAAT	1116
QY	785	ACT---CATGAGTGGAGTGGGAAAGTGTGTGTTAGGATGTCTTCAATTTTCTTCCCTC	841
DB	1117	GTTACCATGGTGGAACTGGCGACGATATTGATCGGAGCATCTCTCTCGCATTCCTC	1176
QY	842	CTTAGCACAAGATCTTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTCAGCAATGGCG	901
DB	1177	CTGGTTGCAAAATACATTTGCAAAAGAATAAGAAACTCTTTTGGTGGCTGCAATTGCA	1236
QY	902	CCATTGACGTCGCTTATATTGGGAAGTCTCTTGGTTTATTTCACTACGCCGAGAGAC	961
DB	1237	CCACTCACTTCGGTGCATCATATCCACCTTTTGTCTATATCACTCGAGCAGCAAGCAT	1296
QY	962	GGTGTGGAAGTATAGGAGAACTGAAGAAGGGTTTTGAATCCACCATCACTCAAAAATCG	1021
DB	1297	GGTGTGCTGATTGTCAAGTACATAAAGAAAGGCATCAATCCACCTTCAGCTAGTCTCATA	1356
QY	1022	GTATTTGTGTCGCTTACATGACTACAGCTGTCAAAACTGGCAATTTGTCTGGCATCATATA	1081
DB	1357	TACTTTCAGTGGCCCAACTGATGAAGAGATTCAGAAATAGGAGTAATAGCTGGCATGATA	1416
QY	1082	TCATTGCGGAAGGAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAATTCATAATTT	1141
DB	1417	GGCCTTAAACAGAAAGCAATTTGCAATAGGAAGAACTTCGCTGGCCTGAAGGACTACAAGAT	1476
QY	1142	GATGCAACAAAGAGATGATAGTATTTCGGACCATGAACGTAGTTGGTTCTTTTCACTCT	1201
DB	1477	GACGAAACAGGAAATGGTGTCTTAGAAACAAATGAACATTTGTTGGTTCAATGACTTCT	1536
QY	1202	TGCTACCTCACAACAGGACCACTTTTCGGGTTTCGGCTGTGAACATAAAGCTGGATGCAAG	1261
DB	1537	TGCTAGCTAGCTACAGGTTCTTCTCGCGGTGAGCAGTCAATTAATCATGCTGGCTGCCAA	1596
QY	1262	ACAGCAGCTTCCAACTTAATATGCTACTTGCAGTAATGTTGACATTTGTTATTCCTGACA	1321
DB	1597	ACAGCGGTATCCAAATATTGTTATGTCAATTTGTAAGTACTGCTCACATCGAGCTGATTA	1656
QY	1322	CCCTGTTTCCATTACACTCCCTGCTGGTGTCTATCAGCTATTTATCGTATCTGCAATGCTT	1381
DB	1657	CCATTGTTTCAAGTATACGCCCAATGCCATCTCTTCCATCATCATATCAGCAGTGTCT	1716
QY	1382	GGACTCATAGATTATGAAGAGCCATCCATCTATTAAAGTTTGACAAATTTGACTTTGTG	1441
DB	1717	GGCTTAGTCGACTACCAACAGCATACCTTATCTGGAAGTTGTATAACTTGAATTTCTG	1776
QY	1442	GTTGTCATGAGTGCAATCAATTTGGGCTGTCTTTGGCAGTGTGAAATTTGGCTTAGTCATA	1501
DB	1777	GCACTGTTAGGAGCAATCTTTGGCGTCATAATTTTTCATCAGTGGAGTATGGTCTGCTAAT	1836
QY	1502	GCTATTGTAATATCTGTACTTTCGGGTACTTCTATTATTATGCAAGCCCAAGGACTTCGTT	1561
DB	1837	GGGTTGGGATCTCATCTCGCTAAAATCTCTTACAAAGTTTACACGACCCAGCAACAGTTTG	1896
QY	1562	TTGGGCAACATTTCAAATTTCTGTGATATACCGAAATTTTGAGCACTATCAAAATCCAAA	1621
DB	1897	CTTGGTAACCTTCCAGAGAACATTAATATAGGAATATCGACCATATCCAGAGCTACC	1956
QY	1622	CATGTTCTCGAATGCTAAATTTCTAGATTTGATGACCAATTTACTTTGCCAATGCCAGC	1681
DB	1957	TTGGTACCAGGAGTGTAAATAGTTAGGTTGAGCTCAGCTACTCTACTTCAAACTCGAAC	2016
QY	1682	TATTTAAGAGAAAGGATCAAGGTTGATGATGAGGAAGAAAGAAATTTAAAGCTACA	1741
DB	2017	TATGTTAAAGACAGAAATCTCAGGTTGGCTGAGGAGCAGGAAGAACGTCAGCAGGAACAG	2076
QY	1742	GGGAGAGCTAGTTTTCAGTATGTTTATTAATTCATATGAGTGTCTTGGAAACATTTGATACA	1801
DB	2077	AAGCTACAGAAAACAGTGTCTGATTTGAGCTATCTCCGTTAATTTGACATTTGATACA	2136
QY	1802	AGTGAATTAAGTATGCTTTGAAGAGGTGAAGAAATTTACAGAGAGAGAGACTACAGCTT	1861
DB	2137	AGTGAATCCATCGCTTGGAGGATCTATTTTAGAGCCTTTGAAAGCGCAAAATCCAGCTG	2196







```
Db 1683 ATGG--GTGGAAGATGAGGATAACCTGTGCAAGAGCGTGGGCAAGATCTGCAATACTT 1739
Qy 1765 TATAATTGATGATGAGTGTCTTTGGAAACATTGATACAAAGTGGGAATAAGTATGCTTGAAGA 1824
Db 1740 GGTCTTGATCTTTGGTGTGTCTCTCTGTCGACAACTCTGGAGTTGGGATGCTACTAGA 1799
Qy 1825 GGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTTGTGTTTGGTCAATCCTGTAAAGTGA 1884
Db 1800 AGTACACAGAGCTCGAACGAAGAGGATCACGATAGCTCTGACGAATCCGAGGCTAGA 1859
Qy 1885 AGTATGAGAACTGAACAAATC---GAAGTCCAAATCATTTAGGGAAGAAATGGAT 1941
Db 1860 GGTAAACAGAGAAAGCTGTGTCTGTGATACGT CAGGGACATCTTAGGGGATGAGTGGGT 1919
Qy 1942 CTATCTGACTGTTGAAGAGGCGGTGGAGCATCAACTTCAATCTACGTGCAAGCAAAAC 2001
Db 1920 CTTCTGACGCTCAAGACGCCATCAGGCGTGTGATACGCGCTCGAGATATCCAGAA 1979
Qy 2002 GAACCCAAAGAAAGATCAA 2020
Db 1980 TRAGGAGAGACGAAGTA 1998
```

## RESULT 12

```
US-10-437-963-73410/c
; Sequence 73410, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 73410
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73694C.1
US-10-437-963-73410
```

```
Query Match 23.3%; Score 571.4; DB 19; Length 2429;
Best Local Similarity 58.1%; Pred. No. 1.3e-142;
Matches 1075; Conservative 0; Mismatches 726; Indels 48; Gaps 2;

Qy 123 TGCACCAAGTCGAGGTTCCACGCCACAGCGCTTTTCAAGTCTCTAAAGTACTCTTTGA 182
Db 2260 TGCACCAAGTGTCTGTCGCGGAGAGGAGGTCTGACGCGGAGGCGCTGAGCAGCGCTGG 2201
Qy 183 AGGAGACTTCTTCTCCCTGATGACCCCTTTGAGGCGAGTTCAAGAACAAAGCCAGCTTCCAAGA 242
Db 2200 CGGAGGTGTTCTTCCCGCAGACCCGCTCCACAGTTCAGAACCAAGTCTGTCGCGCGCG 2141
Qy 243 AGTTCATGCTTGCCCTTCAGTCTTCTTCCCGATTTTGCATGCGGCTCCCAATACACCT 302
Db 2140 GGTGTGTCTGCGCTGCGACTTCTTCCCATCTTCCCATCTTCCCATCTTCCCATCTTCCCAT 2081
Qy 303 TTCAGTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGTAGCTTGGCCATTTCTCTC 362
Db 2080 TCCGCTCTCTCCGCTCCGAC----- 2061
Qy 363 AGGGCATCAGTTATGCAAGCTCGCAACCTCCCTCCAATTTCTTGACTATATTTCCGAGCT 422
```

```
Db 2060 --GGAATCAGCTACGCCAAGCTCGCCAAAGCTCGCCCTCCAAATCATTTGGACTATATTTGAGCT 2003
Qy 423 TTATATACCAACCATTTGATTTATGCGATGATGGTAGCTCGAGGAGATTTGGCAGTGGGACTG 482
Db 2002 TGTGCGCGCGCTGATCTACTCTGTTGGTAGCTCGCGGATCTGGCTGTAGGCGCG 1943
Qy 483 TGGCGGTGGATCGCTTCTGATGGGTTCGATGTTGAGTAATGCCGTGTGATCCCAATGAAG 542
Db 1942 TGTGATAGCTGCTGCTGATGGGTTCGATGCTCGGCAAGCGGTGTGCGCGGACGAG 1883
Qy 543 ACCCAAGCTTTTACCTCCACCTGCTTTCACAGCTACATTTATTTGCTGCTGTTTTCAGS 602
Db 1882 AGCCCATCTCTACCTCCAGCTCGCTTCCCTCCACCTTCTTTCGCGCGGCTTCTTCCAAG 1823
Qy 603 CTGCTTGGGTCTGTTTAGTTGGGTTCGATCGTGAATTTTCTGTACATGCAACCATAA 662
Db 1822 CTTCTCTCGCTTCTCGGCTGGGTTCATCGTGAATTTCTGTGCGAAGCGACGCTGA 1763
Qy 663 TAGGGTTATGGAAGAGCAGCAAGGTGGTGTCTGCGAGCAACTAAATCGATTTCTTG 722
Db 1762 CGGGATTTCATGGCGCGCGCGCATCATGCTGTGCTGCGAGCAGCTCAAGGGATGCTCG 1703
Qy 723 GCCTTGAGCATTTTCAACCATGAGCTGATATCATATCATGATGCTGCTGTTTTCACCC 782
Db 1702 GGATCATCCATTTACGTGCGAGATGGGGTTCGTCCAGGTGATGACATCCGCTTCCAAGC 1643
Qy 783 AAACCTCATGAGTGGAGTGGGAAAGTGTGTGTAGGATGTGTCTTCAATTTCTTCTCTCC 842
Db 1642 ACCACGAGGTGGGCTGGGAGCAATCTCTCATGGGCTCGCTTCTTCTCGCGTCTCTCC 1583
Qy 843 TTAGCACAGATACCTTTCAGCAAAAACGACCAAGGTTTTTTTTGGGTGTCAGCAATGGCGC 902
Db 1582 TCACCACACGCCACATCAGCGCCAGAAACCAAGGCTTCTTGGGTGTGAGCAGCTGCTC 1523
Qy 903 CATTCAGCTCGTTATATTTGGGAAGTCTCTGTTGTTATTTTCTACTCAGCGCGAGACGAG 962
Db 1522 CACTGACATCAGTGTATCTCTCATCATCATCTCGTTCGTTCAGCAAAAGC-----TCATG 1469
Qy 963 GTGTTGAAGTGTAGGAGAACTGGAAGAGGGTTGAAATCCACCATCACTCACAAATCTCG 1022
Db 1468 GCATCAGTGTATTTGGGATCTCCCAAGGAGTTGAACCTCTCTTCAGGGAATATGCTGA 1409
Qy 1023 TATTTGTGCGCTTACATGACTACAGCTGTCAAACTGGCAATGTCTGTGGCATCATAT 1082
Db 1408 CCTTCAGTGGCTCCTACGTAGGACTGGCTTAAACACAGGAATATGACTGTCATAT 1349
Qy 1083 CACTTGGGAGGAATAGCAGTAGGAAGAGCTTTTGCATGTATTAATAATTAACAATTTG 1142
Db 1348 CTCTCACTGAAGGATAGCAGTAGGAGGACATTTTGCATCCATCAACAACTACCGGTTG 1289
Qy 1143 ATGGCAACAAAGAGATGATAGCTATTTGGAGCAATGAAGTGTGTTCTTTCACCTCTT 1202
Db 1288 ATGGAACCAAGGATGATGCGCATTTGTGTATGAACATGGCTGCTCATGTGCTCTCT 1229
Qy 1203 GCTACTCTCAACAGGACCATTTTCGCGTTTCGGCTGTGAATCAATTAACCGCTGGATGCAAGA 1262
Db 1228 GCTATGTCACACAGGATCGTTCTCGAGTTCGCGGTGAATACAGCGCGGGGTGCAAGA 1169
Qy 1263 CAGCAGTTTCAACATTAATATGTCATTTGACATGTTGACATGTTATTTCTGTCGACAC 1322
Db 1168 CGCGGTGTGCAACATCGTGTGAGTGGGTGCGCGGTGCTGCTGCTGCTGCTGCTGCTG 1109
Qy 1323 CCTTGTTCATTAACCTCCCTGTTGCTATCATAGTATTAATCATCTGCAATGCTTTG 1382
Db 1108 CGCTGTTCATTAACCTCCCAAGGATGCTCTCGCGGATCATCATCACCGCGTATCG 1049
Qy 1383 GACTCATAGATTAAGACAGCAGCATCTATTTAAGTTTGACAAATTTGACTTTTGTG 1442
Db 1048 GCCTCATCGAGTCCGGGGGCGCGCAGGCTGTGGAAGGTTCGCAAGCTCGACTTCTCG 989
Qy 1443 TGTGATGAGTGCATACATTTGGCGTGTGTTGGCAGTGTGAAATTTGGCTTAGTATAG 1502
Db 988 CCTGCAATGGCCCTTCTCGCGCTCTCTCTCGTCTCGTCCAGATGGGCTCGCCATCG 929
```

Qy	1503	CTATTGTTAAATATCTGTACTTTCGGGTACTTCTATTATTTATTGCAAGGCCAAGGACATTCGGTTT	1562
Db	928	CCGTCGGCATCTCCCTCTTCAAGATCCCTCTCCAGGTCAACCGCCCAACATGGTGTCA	869
Qy	1563	TGGGCAACATTCGAAATCTCTGTGATATACCGAAATTTGGAGCATATCAAAATGCAAAAC	1622
Db	868	AGGCGGTCTGTCCTCCGCGCAGCGGAGCTACCGAGCATGGCGAGTACAGGGAGGCCATGC	809
Qy	1623	ATGTTCTCGAATGCTAAATCTTAGAGATTGATGCACCAATTTACTTTTGCCAAATGCCAGCT	1682
Db	808	GTGTGCGGTCTTCTCTGCTGCTGGGGTGGAGTCCGGCATCTACTTTCGCCAACTCCATGT	749
Qy	1683	ATTTAAGAGAAAGGATCAACAGGTGGATTGATGAAGAAAGAAAGAAATTAAGCTACAG	1742
Db	748	ACCTCGCGAGAGGATCATGAGTTCTCCGGGAGGAAGACGAGCGGCCGCCAAGTGCA	689
Qy	1743	GGGAGACTAGTTTGCAGTATGTTATAATGATATGAGTGTGTTGGAAACATTTGATACAA	1802
Db	688	ACCAAGTCCCTGTGCAGATGCATCATCTCGACATGAGTGTGTTGCAGCGATGCACAGA	629
Qy	1803	GTGGAATAGTATGCTTGAAGAGGTGGAAGAATTTACAGAGAGAGAGAGCTACAGCTTG	1862
Db	628	GTGGCCTCGATGTCATTTGCAGAGCTGAAGAAGGTGCTGGAAAAAGAAACATCGAGCTTG	569
Qy	1863	TTTTTGTCAATCTGTAAGTGAAGTGATGAAGAAATCTGAACAAATTCGAAAGTTTCCAAAATC	1922
Db	568	TGCTGGCCAAACCGGTTTGATCGGTGACCGAGAGGCTGTACAACTCGGTGTGGCGAAGA	509
Qy	1923	ATTTAGGGAAGAAATGGATCTATCTGACTGTGTTGAAGAGGCCGTTGGAGC	1971
Db	508	CGTTTGGCTCAGACCCGCGTGTCTTTCAGCGTCCGCGAGCGCGTCGCGC	460
RESULT 13			
US-10-425-114-15719			
; Sequence 15719, Application US/10425114			
; Publication No. US20040034888A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Jingdong			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Screen, Steven E			
; APPLICANT: Tabaska, Jack E			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(53313)B			
; CURRENT APPLICATION NUMBER: US/10/425,114			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 73128			
; SEQ ID NO 15719			
; LENGTH: 2656			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; OTHER INFORMATION: Clone ID: LIB3060-059-F10_FLI			
US-10-425-114-15719			
Query Match 23.2%; Score 567; DB 18; Length 2656;			
Best Local Similarity 57.0%; Pred. No. 2.1e-141;			
Matches 1060; Conservative 0; Mismatches 795; Indels 6; Gaps 1;			
Qy	123	TGCACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTGA	182
Db	300	TGCACCAAGGTGTCTGCTGCCGGAGCGCGACGACGCGAAGCGCTCGGCGAGCGCCTGG	359
Qy	183	AGGAGACTTCTTCCCTGATGACCTTTTGAGGCGAGTTCAAGAACAGCCAGCTTCCAGA	242
Db	360	CCGAGGTGTTCTTCCCGAGACCCGCTGCACAGTTCAAGAACACAGTCGTCTGGCGGCG	419
Qy	243	AGTTTCATGCTTGGCCTTCAGTCTCTTCTCCCAATTTTGGAAATGGGCTCCCAAAATACACCT	302

Qy	1383	GACTCATAGATTATGAACGACGCATCCATCTATTAAAGTTTGAACAAATTTGACTTTGTGG	1442
Db	1554	GGCTGGTGGAGCTGCGCGCGCCGCGCAGGCTGTGGAAAGTGGACAAGCTGGACTTCTCTGG	1613
Qy	1443	TGTGCATGAGTGCATACATTTGGCGTGGTCTTTTGGCAGTGTGTGAATTTGGCTTTAGTCATAG	1502
Db	1614	CGTGGTGGCGCGCTTCTCGCGGTGCTGCTGGTCTCGTCGACAGCGGCCCTGGCGGCTCG	1673
Qy	1503	CTATTGTAAATATCTGTACTCTCGGGTACTTCTATTATTTTGAAGGCCAAGGACATTCGTTT	1562
Db	1674	CCGTCGGCATTCTGCTCTTCAAGGTCCTGTCTGACGGTCAACCGGCCCAACGCTCGTGGTGG	1733
Qy	1563	TGGGCAACATTCCAAATCTCTGTGATATACCGAAATGTTTGACACTATCAAAATGCAAAAC	1622
Db	1734	AGGGCTCGTCCGGGGACGCAGAGCTACCGCAGCGCTGGCGCAGTACCGCGAGGCCGTCC	1793
Qy	1623	ATGTTCTTGGAAATGCTAATCTTAGAGATGTGATGCAACCAATTTATCTTTCGCAATTCGCACT	1682
Db	1794	CGCTGCCGGGCTTCTCTCGCTCGCGCGTCTGAGTCCGCGCTCTACTCTCGCCCACTCCATGT	1853
Qy	1683	ATTTAAGAGAAGGATCAACAGGTGGATTTGATGAAGAAGAGAAGAAATTAAGACTACAG	1742
Db	1854	ACCTGGTGACGGGTCTATGCGTACTCTCCGGCAGCAGGAGGCGCGGCTCAAGTCCA	1913
Qy	1743	GGGAGACTAGTTTTCAGTATGTTATTAATTTGATATGAGTGTCTGTTGGAAACATTTGATACAA	1802
Db	1914	ACCACCCCTCCATCCGATCGCTCGTCTCGCATGGGCGCGCTCGCGCGATTCGACACGA	1973
Qy	1803	GTGGAAATAGTATGCTTTGAAGGCTGAAGAGATTTACAGAGAGAGAGAGCTACAGCTTG	1862
Db	1974	GCGGTCTTAGCGGCTGTCCGAGCTCAAGAAAGTCTCTGCAACAAAGAAACATCTGHAGCTGG	2033
Qy	1863	TTTTTGGTCAATCTGTAAAGTGAAGTGAAGAAACTGAACAAATTCGAAGTTTCCAAATC	1922
Db	2034	TGCTTGCCAAACCGGTGGGGTCCGTTGGCGGAGAGGATGTTTCAACTCGCGCGTGGGCGGAGA	2093
Qy	1923	ATTTTAGGGAAGAAATGGATCTATCTGACTGTTTGAAGAGCGCGTTGAGCATGCAACTTCA	1982
Db	2094	GCTTCGGGTCGGGCGCCTCTTCTTCAGGTTAGGGAGGCCGTTCGCGGGGGGGCGTGCA	2153
Qy	1983	A	1983
Db	2154	A	2154

```

RESULT 14
US-10-425-115-75763
; Sequence 75763, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 75763
; LENGTH: 2713
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_169123C.1
US-10-425-115-75763

```

Query Match 23.2%; Score 567; DB 20; Length 2713;  
Best Local Similarity 57.0%; Pred. No. 2.1e-141;  
Matches 1060; Conservative 0; Mismatches 795; Indels 6; Gaps 1;

Qy	123	TGACCAAGTGTGAGTTCCACCGCCACAGCCGTTTTTCAAGTCTCTAAAGTACTCTTTTGA	182
Db	357	TGACCAAGTGTGTGCTGTCGGAGCGCGGACGACGCGAAAGGCGCTGCGGACAGCGCCTGG	416
Qy	183	AGGAGACTTTCTTCCCTGATGACCCCTTTTGGGCAGTTTCAAGAAACAAGCCAGCTTCCAAGA	242
Db	417	CCGAGGTGTTCTTCCCGGACGACCCGCTGCACAGTTTCAAGAACCAAGTCTGTGCGCGCGC	476
Qy	243	AGTTCAATGCTTGGCCTTCAAGTTCTTCTTCCCAATTTTTCGAATGGGTCTCCAAATACACCT	302
Db	477	GCCTCGTGTGGCGCTGCACACTTCTTCTCCCATCTTCCAGTGGGGTCTCGCCTACAGCC	536
Qy	303	TTCAAGTCTTGAAGCTGACCTCATAGCTGGGATCACCATCGCTAGCTTGGCCATTTCCCTC	362
Db	537	CGGCGCTCTCGCGCTCCGAGCTGTGCGCGGCTCACCATTTGCAGCGCTCGCCATCCCGC	596
Qy	363	AGGCATCAAGTATGCCAGCTCGCAACCTCCCTCCAATTTTGGACTATATTCAGACT	422
Db	597	AGGAATCAAGTACGCCAGCTCGCCAACTGCGCGCCATCTGTGGCCTATATTCAGCT	656
Qy	423	TTATACCAATTTGATTTATGCGATGATGGGTAGCTCGAGGGAATTTGGCAGTGGGACTG	482
Db	657	TCGTGCGCGCTCATCTACGCGCTGTGCGGAGCTCGCGGACCTGGCGTGGGCGCGG	716
Qy	483	TGGCGGTTGATCGCTTCTGATGGGTTGATGTGAGTATGCGGTGATCCCAATGAAG	542
Db	717	TGTCATCGCTCGCTGTGATGGGGTCCATGTCCGGGACGCGCTGTCCGCGGACGAGC	776
Qy	543	ACCCAAAGCTTACCTCCACCTCGCTTTCACAGCTACATTTATTTGCTGGTGTGTTTTCAGG	602
Db	777	AGCGCTCTCTACCTGCGACTGGCCTTCAACGCACTTCTTTCGCGGCTCTTCCAGG	836
Qy	603	CTGCGTTGGGTCGTTTAGGTTGGGTTGATCGTGGATTTTCTGTCACTGCAACCATAA	662
Db	837	CGTCCCTGGGATTCCTCAGGCTGGCTTCACTGTGGACTTCTGTCCAAGCGCAGCGCTGA	896
Qy	663	TAGGTTTCATGGAGAGACGACAGGTGTGTGTCTGAGAGAACTAAATACGATTTTG	722
Db	897	CGGCTTTCATGGGCGGCGCGCTCATCTGTGTCTGTGACAGAGCTCAAGGGGCTGTCTCG	956
Qy	723	GCCTTGAGCATTTACCCATGGAGCTGATATCATATCAGTGATGCGCTCTGTGTTTTCACCC	782
Db	957	GCATCTCCCACTTCACTCCCAATGGGATTTCTCGACGTCATGCGCTCGTGTCAACC	1016
Qy	783	AAACTCATAGTGGAGTGGGAAAGTCTGTGTTAGGATGTCTTCAATTTTCTTCTCTCC	842
Db	1017	GCCACGAGTGGAGTGGACAGCATCTCATGGCTCCGCTTCTCGCCATCTCTCC	1076
Qy	843	TTAGCAAGATACTTTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTGAGCAATGGCGC	902
Db	1077	TCCTCAGCGCCAAATCAGCGCCAGGAAACCAAGCTTTTTCTGGGTATCAGCAGGTGCTC	1136
Qy	903	CATTGACGTCGGTTATTTGGGAAGTCTCTGTGTTTATTTCACTCAGCGCGAGAGCAGC	962
Db	1137	CCCTGGCGTCGGTGATCATCTCCACCATCTCTCTTTCATCTGGAAATCCCCAGACATCA	1196
Qy	963	GTGTTGAAGTGTAGGAGAACTCAAGAAGGGTTTTGAATCCACCATCACTCACAAAATCTGG	1022
Db	1197	GTGTTT-----ATTGGCATCTCCCAAGGGAGTGAACCTCTTTCGGCGGACATGCTCA	1250
Qy	1023	TATTTGTGTGCGCTTACATGACTACAGCTGTCAAAACTGGCATTTGTCTGTGGCATCATAT	1082
Db	1251	GCTTCAGCGCTCTCATGTGCGCTGACGATCAAAACCGGATCATGACAGGCATCTCTGT	1310
Qy	1083	CACTTGGGAAGGATAGCAGTAGGAGAGAGCTTTTGAATGTATATAAATTACAAATTTG	1142
Db	1311	CTTTAAACAGAAAGGATCGCATGTGGGAGGACCTTTCGCGTCCATCAACAACATCAAGGTG	1370
Qy	1143	ATGCAACAAGAGATGATAGCTATTGGGACCATGAACGATGTGTTGTTTCTTTCACCTCTT	1202
Db	1371	ACGGGAACAAGGAGATGATGCGCATCGGCTGTATGAACATGCGGGGCTCTCGGCTCTCT	1430
Qy	1203	GCTACCTTCAACAACAGGACCAATTTTCGGGTTTCGGCTGTGAACATATAACGCTGGAATGA	1262

Db 1431 GCTACGTGACGCGGGTCTTCTCCGGTCCGGGTGAACATACAGCGGGGCTCAGGA 1490  
Qy 1263 CAGCAGCTTCCAAACATTATAATGTACATTGCAAGTAAATGTTGACATGTTATTTCTTGACAC 1322  
Db 1491 CGGCGCTGTCCAACTGTCGTGATGGCGGCGGTGCTGGTGAAGCTGCTGTTCTCATGC 1550  
Qy 1323 CTTTGTTCATTAACATCTCCCTGGTGGTGTATCAGCTATTAATGCTATTAATGCAATGCTTG 1382  
Db 1551 CGCTGTTCCATACACCCCGAACGTCATCTGGCGCGCATCATCATCAGCGGCGGTGG 1610  
Qy 1383 GACTCATAGATTATGACGAGCCATCTATTTAAAGTTGACAAATTTGACTTTTGG 1442  
Db 1611 GGCTGGTGACGTGGCGGCGCGGCGAGCTGTGGAAGGTGCAAGCTGGHACTTCTCTGG 1670  
Qy 1443 TGTGATGAGTGATACATATGCGGTGGTCTTTGGCAGTGTGCAAAATTTGGCTTAGTCATAG 1502  
Db 1671 CGTGGTGGCGGGTTCCTCGGGTGTCTGCTGTGTCGAGACGGGCTGGGCGTGG 1730  
Qy 1503 CTATTGTAATATCTGTACTTCCGGGTACTTCTATTTATTGCAAGGCCAAGGACATTCGTTT 1562  
Db 1731 CGGTGGCATCTCGCTCTTCAAGGTCTGCTGCAGGTCAACCGCCCAACGTCGTGGTGG 1790  
Qy 1563 TGGGCNAACATTCGTAATCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAAC 1622  
Db 1791 AGGGCTTCGTCCCGGGGACGACAGACTACCGCAGCGTGGCGCAGTACCGGAGGCGCGTCC 1850  
Qy 1623 ATGTTCTCTGGAATGCTAAATTTCTAGAGATTGATGCACCAATTTACTTTTGCCAAATGCCAGT 1682  
Db 1851 GCGTGGCGGGCTTCTCTGTCGTGGGGTGCAGTCCGCCGTCTACTTGGCCAACTCCATGT 1910  
Qy 1683 ATTTAAGAGAAAGGATCAACAGGTGGATTGATGAAGAAAGAAAGAAATTAAGCTACAG 1742  
Db 1911 ACTGTGTGAGCGGGTCATGCGCTACCTCCGCGACGAGGAGGCGCGCTCAAAGTCCA 1970  
Qy 1743 GGGAGACTAGTTGCGAGTATGTTAATTTGATGATGAGTCTCTTGGAAACATTGATACAA 1802  
Db 1971 ACCACCCCTCCATCCGATCGCTCGTCTCGACATGGGCGCCGTGCGGCGATTCGACAGA 2030  
Qy 1803 GTGGAAATAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGCTACAGCTTG 1862  
Db 2031 GCGGTCTAGACGGCTGTCCGAGCTCAAGAAAGTCTCGACAAAGAAACATCGAGCTGG 2090  
Qy 1863 TTTTGTCTAATCTGTGAAGTGAAGTGAAGAACTGAAGAACTGAAGTTCCAAATC 1922  
Db 2091 TGCTTGCCAAACCGGTGGGGTGGTGGCGAGAGGATGTTCAACTCGGCGGTGGCGAGA 2150  
Qy 1923 ATTTAGGGAAGAAATGGATCTATCTGACTGTTGAGAGGCCGCTTGGAGCATCAACTTCA 1982  
Db 2151 GCTTCGGGTGGGCGCGCTCTTCTTCAGGTAGACGAGGCCGTGCGGCGGGGGCGTGCA 2210  
Qy 1983 A 1983  
Db 2211 A 2211

RESULT 15  
US-10-424-599-32308  
; Sequence 32308, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 32308  
; LENGTH: 536

; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_129179C.1  
US-10-424-599-32308  
Query Match 21.2%; Score 520.2; DB 18; Length 536;  
Best Local Similarity 98.5%; Pred. No. 3.3e-129; Indels 0; Gaps 0;  
Matches 525; Conservative 0; Mismatches 8;  
Qy 7 GCTAGCTCGCACATTAAGTTATATAACACATATTTGCTTGTCTTAGAAATACTATTATTGA 66  
Db 4 GCTAGCTCGCACATTAAGTTATATAACACATATTTGCTTGTCTTAGAAATACTATTATTGA 63  
Qy 67 AGATATGGGAGTGTAGATTATGATACCCCTTTGGGCATGAACAACCTTTGAGAGATGCA 126  
Db 64 AGATATGGGAGTGTAGATTATGATACCCCTTTGGGCATGAACAACCTTTGAGAGATGCA 123  
Qy 127 CCAAGTCGAGGTCCACCGCCACAGCGTCTTTCAAGTCTCTAAAGTACTCTTTGAAGGA 186  
Db 124 CCAAGTCGAGGTCCACCGCCACAGCGTCTTTCAAGTCTCTAAAGTACTCTTTGAAGGA 183  
Qy 187 GACTTTCTTCCCTGATGACCCCTTTGAGGAGCTTCAAGAACAGCCAGCTTCCAAGAGTT 246  
Db 184 GACTTTCTTCCCTGATGACCCCTTTGAGGAGCTTCAAGAACAGCCAGCTTCCAAGAGTT 243  
Qy 247 CATGTTGGCCTTCAGTTCCTTCTCCCATTTTCGAATGGGCTCCCAAAATACACCTTTCA 306  
Db 244 CATGTTGGCCTTCAGTTCCTTCTCCCATTTTCGAATGGGCTCCCAAAATACACCTTTCA 303  
Qy 307 GTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCTCAGGG 366  
Db 304 GTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCTCAGGG 363  
Qy 367 CATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTGGACTATATTCGAGCTTTAT 426  
Db 364 CATCAGTTATGCCAAGCTCGCAACCTCCCTCCAAATCTTGGACTATATTCGAGCTTTAT 423  
Qy 427 ACCACCATTTGATGCGATGATGGTAGCTCGAGGATTTGGCAGTGGGAGCTGTGGC 486  
Db 424 ACCACCATTTGATGCGATGATGGTAGCTCGAGGATTTGGCAGTGGGAGCTGTGGC 483  
Qy 487 GGTGGATCGCTTCTGATGGGTTCGATGTCAGTAATGCCGTTCATCCCAATG 539  
Db 484 GGTGGATCGCTTCTGATGGGTTCGATGTCAGTAATGCCGTTCATCCCAATG 536

Search completed: September 12, 2005, 14:59:41  
Job time : 5645 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2005, 11:07:23 ; Search time 57 Seconds  
(without alignments)  
890.550 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGWNV 680

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3475	100.0	680	4	US-09-720-317A-18
2	2642.5	76.0	658	4	US-09-720-317A-24
3	2575.5	74.1	646	4	US-09-720-317A-26
4	2314	66.6	646	4	US-09-720-317A-28
5	2164	62.3	579	4	US-09-720-317A-4
6	1901	54.7	621	4	US-09-720-317A-16
7	1890.5	54.4	631	4	US-09-720-317A-29
8	1843.5	53.1	688	4	US-09-720-317A-2
9	1798.5	51.8	656	4	US-09-720-317A-20
10	1760.5	50.7	660	4	US-09-720-317A-30
11	1723.5	49.6	660	4	US-09-720-317A-23
12	1503	43.3	644	4	US-09-720-317A-25
13	997	27.7	685	4	US-09-720-317A-31
14	961.5	27.1	593	4	US-09-720-317A-22
15	731.5	21.1	226	4	US-09-720-317A-8
16	687.5	19.8	233	4	US-09-720-317A-27
17	687.5	19.8	596	4	US-09-252-991A-23812
18	675.5	19.4	311	4	US-09-720-317A-6
19	650.5	18.7	744	4	US-09-785-381-3
20	650	18.7	744	4	US-09-785-381-1
21	643.5	18.5	590	4	US-09-902-540-14944
22	641.5	18.5	748	4	US-09-949-016-10387
23	612	17.6	565	4	US-09-602-787A-616
24	587.5	16.9	780	4	US-09-785-381-11
25	586.5	16.9	764	1	US-08-424-567-2
26	586.5	16.9	764	2	US-08-711-928-2
27	586.5	16.9	764	3	US-09-184-937-2

28	586.5	16.9	790	4	US-09-949-016-11220	Sequence 11220, A
29	576.5	16.6	803	4	US-09-949-016-11498	Sequence 11498, A
30	567	16.3	828	4	US-09-248-796A-20746	Sequence 20746, A
31	549.5	15.8	663	4	US-09-875-811-6	Sequence 6, Appli
32	549	15.8	656	4	US-09-875-811-10	Sequence 10, Appl
33	546.5	15.7	679	4	US-09-875-811-2	Sequence 2, Appli
34	489	14.1	616	4	US-09-543-681A-4421	Sequence 8, Appli
35	485.5	14.0	605	4	US-09-875-811-8	Sequence 12, Appl
36	485	14.0	598	4	US-09-875-811-12	Sequence 4, Appli
37	482.5	13.9	621	4	US-09-875-811-4	Sequence 7, Appli
38	454	13.1	970	4	US-09-795-927-7	Sequence 21805, A
39	449	12.9	535	4	US-09-252-991A-21805	Sequence 13374, A
40	426.5	12.3	472	4	US-09-902-540-13374	Sequence 1015, Ap
41	402.5	11.6	575	4	US-09-438-185A-1015	Sequence 14, Appl
42	381	11.0	159	4	US-09-720-317A-14	Sequence 4544, Ap
43	369	10.6	566	4	US-09-543-681A-4544	Sequence 5244, Ap
44	363	10.4	533	4	US-09-107-532A-5244	Sequence 10, Appl
45	360.5	10.4	147	4	US-09-720-317A-10	

#### ALIGNMENTS

RESULT 1

US-09-720-317A-18

; Sequence 18, Application US/09720317A

; Patent No. 6696292

; GENERAL INFORMATION:

; APPLICANT: Stephen M. Allen

; APPLICANT: Saverio C. Falco

; APPLICANT: Catherine J. Thorpe

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167

; CURRENT APPLICATION NUMBER: US/09/720,317A

; PRIORITY FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 18

; TYPE: PRT

; LENGTH: 680

; ORGANISM: Glycine max

US-09-720-317A-18

Query Match 100.0%; Score 3475; DB 4; Length 680;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HELARTLSYTHICLLRNTIIEDMGSDVDYEYPLGMNPNFVHVEVPPQPPFKSLKYSL 60

Db 1 HELARTLSYTHICLLRNTIIEDMGSDVDYEYPLGMNPNFVHVEVPPQPPFKSLKYSL 60

QY 61 KETFPDDPLRQKPKASKKFMGLQPPFPIFEWAPKYTFQFLKADLIAGITIASLAIP 120

Db 61 KETFPDDPLRQKPKASKKFMGLQPPFPIFEWAPKYTFQFLKADLIAGITIASLAIP 120

QY 121 QGISYAKLANLPPIILGLYSSFTPLIYAMGSSRDLAGTAVGSLMGSMLSNVDPNE 180

Db 121 QGISYAKLANLPPIILGLYSSFTPLIYAMGSSRDLAGTAVGSLMGSMLSNVDPNE 180

QY 181 DPKYLHLAFTATLFGVFOALGFRGLIVDFLSHATIIIGFMGAATVVCQLQKLSIL 240

Db 181 DPKYLHLAFTATLFGVFOALGFRGLIVDFLSHATIIIGFMGAATVVCQLQKLSIL 240

QY 241 GLEHFTHGADIIISVMSVFTQTHWRWSAVLGCVFIFFLSTRYFSKCRPRFFWVSAMA 300

Db 241 GLEHFTHGADIIISVMSVFTQTHWRWSAVLGCVFIFFLSTRYFSKCRPRFFWVSAMA 300

QY 301 PLTSVILGSLLYFTHAEKHGVEVIGELKKGLNPPSLTNLNVFVSPVMTTAVKTGIVGII 360

Db 301 PLTSVILGSLLYFTHAEKHGVEVIGELKKGLNPPSLTNLNVFVSPVMTTAVKTGIVGII 360















RESULT 12  
US-09-720-317A-25  
; Sequence 25, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 25  
; LENGTH: 644  
; TYPE: PRT  
; ORGANISM: Stylosanthes hamata  
US-09-720-317A-25

Query Match 43.3%; Score 1503; DB 4; Length 644;  
Best Local Similarity 47.0%; Pred. No. 1.2e-142;  
Matches 301; Conservative 124; Mismatches 193; Indels 22; Gaps 6;

QY 33 LGNNFTRVHQ--VEVPPQPFKSLKYSIKETFFDDPLRQFNKPKASKMGLQF-- 88  
DB 4 LGTEQFSRSQWLNPNPPLTKFLGLKD-----NKFTSSSKKTRAVSFLA 55

QY 89 -FPPIFEWAPKYTFQFKADLIAGITIASIAIPOGISYAKLANLPILGILYSSFIPLIY 147  
DB 56 SLEPILSWRTYSATFKDGLSLASLSIQSGYANLAKLDPOYGLYTSVIPVIY 115

QY 148 AMWGSSRDLAGVTAVGSLMGLMNAVDNPDPKLYHLAFTATLFAVFOAALGLPR 207  
DB 116 ALMGSSREIAIGPVAVVMSSSLVPKVIDPDADPNDRNLVFTVTLFAGIFQAFGLR 175

QY 208 LGLIIVPLSHATIIIGWGNATVVCLOOLKSIILGLEHFTGADIISVMSVFTQTHE 264  
DB 176 LGLVDFLSHAALVGMAGAAIIVIGLQOLKGLGLTHFTTKTDAVALKSVYTSLHQQT 235

QY 265 ---WRWESAVLCGVFIFFLLSRYFSKRPFFWVSAMAPLTSVILGSLVYVTHAEKH 320  
DB 236 SSENWSPLEFVICCSFLIFLAARFGRNKKFENLPAIAPLSSVILSVLFLSKGDH 295

QY 321 GVEVIGELKKGLNPPSLTNLTVFVSPLYMTTAVKTIIVVGIISLAEGIAVGRSPFMYKYN 380  
DB 296 GVNIKHVQGLNPSSVHKQLQNGPHVGOAAKIGLISAIITAEIATVGRSPANIKGYHL 355

QY 381 DGNKEMIAIGTMNVGVSFTSCYLLTTPFSSAVNAGCKTAASNILMSLAVMLTLLPLT 440  
DB 356 DGNKEMLAGCMNAGSLTSYCYSTGTSFSTAVNFSAGCKTAVSNVMAVTLVLLCLEFT 415

QY 441 PLPHYTPLVLSAIIYSAMGLDYEAHLFKVDKDFVFCMSAVIGVVGVSVEIGLVI 500  
DB 416 RLITYTPMAILASIIISALPGLIDIGEAHYHVKVDKDFLACLAFGLVFSIEIGLII 475

QY 501 AIVISLVRLLFTFTRPTFVLGNIPNSVIRYRNVEHYQNAKHVPGMLILBIDA-PIYFANA 559  
DB 476 ALSISPAKILLOAIRPGVEVLGRIPTEAYCDVAQPMVAVTTPGILVIRISSGSLCFANA 535

QY 560 SYLRERITRWI-DEBERIKATGETSLOVVIDMSAVGNIDTSGISMLSEVKKITERREL 618  
DB 536 GFVREIRLKWVEDEEQNIEAAKGRVQAIIDMTDLTNVDTSGILALBELHKKLLSRGV 595

QY 619 QLVLVNVPVSVMKLNKSKFQHLKKWIIYLTVEEAVGAC 658  
DB 596 ELAWNPRVEVHKLKVFVVDKIGKERVFLTVAEAVDAC 635

RESULT 13

US-09-720-317A-31  
; Sequence 31, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 31  
; LENGTH: 685  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-720-317A-31

Query Match 28.7%; Score 997; DB 4; Length 685;  
Best Local Similarity 34.0%; Pred. No. 1.8e-91;  
Matches 217; Conservative 140; Mismatches 232; Indels 50; Gaps 13;

QY 47 PPOQPFKSLKYSIKETFFDDPLRQFNKPKASKMGLQFPPPIFEWAPKYTF-QFLK 105  
DB 51 PPSIIF-----DIFSGWTAKIKRMRLVDWIDTLFPCFRVIRTYRWSEYFK 96

QY 106 ADLIAGITIASIAIPOGISYAKLANLPILGILYSSFIPLIYAMWGSSRDLAGVTAVGS 165  
DB 97 LDLMAGITVGMILVPMQMSYAKLAGLPIYGLYSSFPVFPVYVAFGSSRQLAIGPVALVS 156

QY 166 LLMGSMLSNVDNPDPKLYHLAFTATLFAVFOAALGLFRLGLIVDFLSHATIIIGFMG 225  
DB 157 LLVSNALGGIADTNBEE--LHIELAILLALLVGLICIMGLLGLMLIRFISHVISGFTS 214

QY 226 GAATVVCLOOLKSIILGLEHFTGADIISVMSVFTQTHEWRWESAVLCGVFIFFLLSTRY 285  
DB 215 ASAVIGLSQIKYFLGYS-IARSSKIVPIVESIAGADKQFQPPFVWGSLILVLQVMKH 273

QY 286 FSKGRPRFFWVSAMAPLTSVILGSLVYVTHAEKHGVEVIGELKKGLN----PPSLTNLV 341  
DB 274 VGKAKKELQFLRAAAPLGTGIVLGTITIAKVHPP--SISLVGEIPIQGLPTFSFSPSDHAK 331

QY 342 FVSPYMTTAVKTIIVVGIISLAEGIAVGRSPFMYKYNINDGNKEMIAICTMNVGVSFTSC 401  
DB 332 TLLP--TSALITG-----VPILESVGIKALAAKNRYELDSNDDLPGLVANILGSLFSA 384

QY 402 YLTTPFSSRSVAVNAGCKTAASNILMSLAVMLTLLPLTFLPHYTPLVVLSAIIYSAMLG 461  
DB 385 YPATGSFSSAVNNSSEAKTGLSGLITGIIIGCSLLFTPMFKYIPOCALAAIIVISAVSG 444

QY 462 LDYDEAAHLFKVDKDFVW-CMSAYIGVVGVSVEIGLVIIVISLVRLLFTFTRPTFV 520  
DB 445 LDYDEAFLMRVDRDFSLMTITSTITLFFG-IEIGVLVGVGSLAFVHESANPHIAV 503

QY 521 LGNIPNSVIRYRNVEHYQNAKHVPGMLILBIDAPIYFANASYLRERITRW---IDE----- 572  
DB 504 LGRLPCTTVYRNKQYPEAYTYNGIVIVRIDSPYFANISYIKDRUREVEAVDKYTRNG 563

QY 573 -DEBERIKATGETSLOVVIDMSAVGNIDTSGISMLSEVKKITERRELQVLVNPVSEVMK 631  
DB 564 LEVDRI-----NFVILEMSPVTHIDSSAVALKELQYKYKTRDIQILAISNPKNVDHL 615

QY 632 KLNKSKFQHLKKWIIYLTVEEAVGACNPNLRASKTNPK 670  
DB 616 TIARSGMVELVGKFNFFVRVHDVAVQVCLQYVQSSNLEDK 654

RESULT 14  
US-09-720-317A-22  
; Sequence 22, Application US/09720317A

```
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-720-317A-22

Query Match      27.7%; Score 961.5; DB 4; Length 593;
Best Local Similarity 35.0%; Pred. No. 5.5e-88;
Matches 204; Conservative 127; Mismatches 235; Indels 17; Gaps 7;

QY  91  PIFEWAPKTF-QFLKADLIAGTIIASLAIPQGISYAKLANLPPTILGLYSSFIPLLIYAM 149
Db  22  PCLAMNRSRWKEDFOADLAAGITGVMLVPQAMSYAKLAGLHPHLYGLTVGVPLFVYAI 81

QY  150  MGSSRDLAVGTVAVGSLLMGMSLSNAVDNEDPKLYLHLAFTATLPAFGVQAAALGIFRIG 209
Db  82  FGSSRLGVPVALVSLLSNVLGIV--NSSSELYTELAIIAFAFWGILECLMALLRLG 139

QY  210  LIVDFLSHATITGFMGGAATVVCLOQLKSLILGLEHTHGDIIISVNRSVFTQTHEWRWES 269
Db  140  WLIRFTSHSVISGFTTASAIIVIGLSQIKYFLGYS-VTRSSKIIPLETIESIAGIDQFSWPP 198

QY  270  AVLGCVFIFPLISTRVFSKRRPFFWVSAMAPLTSVILGSLAVYFTHAEKHGVEVIGELK 329
Db  199  FWGSAFLVILLIMKGLGNKRLFRASGPLTAVVLGTLFVKIPRPT--AISVVGEIP 256

QY  330  KGLNPPLTNLNVFSP-----YMTAVKTGIVVGIISLAEGIAVGRSFAMYNKYNIDGNKE 385
Db  257  QGL--PS-----PSIPRGFEHLMSLPTAILITGVAILSVGIKALAAKNGYELDSNKE 309

QY  386  MTAIGTMNVVGSFTCYLTGTFPSRGAANNAGCKTAAANSIIMSLAVMLTLLPLTPLPHY 445
Db  310  LFLGLSLNICGSPFSAYPATGSPFSRGAHVHESAKTGLSGIIMGIIICSALLFMTPLFTD 369

QY  446  TPLVLSAIVSAMGLIDYEAHLFKVDKDFVVCMSAYIGVWFGSVGEIGVIAIVIS 505
Db  370  IPQCALAAIVISAVTGLVDYEEAIFLWGDIDKDDFLWAMTFTTTLTFGIEIGVLGVGFS 429

QY  506  VLRLVLLFIARPTFVLGNIPNSVIYRNVEHYQNAKHVPGLMLILEIDAPIYFANASYLRER 565
Db  430  LAFVHESANPHAVLGRLPGTIVYNTLOYEAYTYNGIVVVRVDADAPIYFANISYIKDR 489

QY  566  ITRWIDEEERIKATGETSLQYVIIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNP 625
Db  490  LREYELKLPNSNRGPDVGRVYFVILEMSPVTYIDSSAVQALKDLHQQEYKARDIQIAIAP 549

QY  626  VSEVMKKLNKSKFQNHGKKWLYLTVEEAAGACNFNLRASKTN 668
Db  550  NRQVHLLLSRAGIIDMIGAWCFVRVHDVAQVCLQHVSRSSSN 592
```

```
RESULT 15
US-09-720-317A-8
; Sequence 8, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
```

```
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helianthus tuberosus
US-09-720-317A-8

Query Match      21.1%; Score 731.5; DB 4; Length 226;
Best Local Similarity 65.6%; Pred. No. 2e-65;
Matches 145; Conservative 31; Mismatches 38; Indels 7; Gaps 1;

QY  467  RAIHLFKVDKDFVVCMSAYIGVWFGSVGEIGVIAIVISVLRLVLLFIARPTFVLGNIPN 526
Db  6  AAHILWTLDDKDFVVCMSAYIGVWFGSVGEIGVIAIVALSURVLLFVSRPRTSTLGLIPD 65

QY  527  SVIYRNVEHYQNAKHVPGLMLILEIDAPIYFANASYLRERITRWIDEEERIKATGETSIQ 586
Db  66  STIYRSMDOYQNAKSVPGILLILQIEAPIYFANSSYLRRERIVRWVDEEDRLKSLKENDLQ 125

QY  587  VYIIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKKLNKSKFQNHGKKW 646
Db  126  VYILALSAVGNIDTSGITMLGEVKKVMERRGLKLVLANPGGGEVIKKMNKAKLIEVIGQBW 185

QY  647  IYLTVEEAAGACNFNLRASKTNPKK-----DETEGWNV 680
Db  186  IYLTVEEAAGACNFMHLHYKNAEKPTSGSESGKSRNDNV 226

Search completed: September 12, 2005, 18:43:11
Job time : 59 secs
```

Result No.	Query			ID		Description
	Score	Match	Length	DB	ID	
1	2330.5	67.1	657	16	US-10-437-963-140667	Sequence
2	2015	58.0	414	15	US-10-424-598-273958	Sequence
3	1853.5	53.3	695	16	US-10-425-115-235264	Sequence
4	1850	53.2	659	15	US-10-424-599-207945	Sequence
5	1811.5	52.1	666	16	US-10-437-963-198336	Sequence
6	1773.5	51.0	662	16	US-10-437-963-179488	Sequence
7	1762	50.7	681	16	US-10-425-115-260436	Sequence
8	1733.5	49.9	656	16	US-10-437-963-175893	Sequence
9	1724	49.6	397	15	US-10-424-598-273959	Sequence
10	1723.5	49.6	653	16	US-10-437-963-179495	Sequence
11	1722.5	49.6	662	16	US-10-425-115-366870	Sequence

Db 18 RVPMPAAKPFLETLGGNM

18 BYDMPBAKPEIETI GGNMKET

[illegible]

18 RVPMPAAKPELETTLGGNMKETFLPDDPFRVRRERCGGRRAAAALRYVPFMEWAPSYTL 77

DO NOT WRITE IN THESE SPACES

QY 102 QFLKADLIAGITIASIAIPOGISYAKLANLPPIILGLYSFIPPLIYAMGSSRDLAGTV 161  
Db 78 GTLKSDLIAGITIASIAIPOGISYAKLANLPPIILGLYSFIPPLIYAMGSSRDLAGTV 137  
QY 162 AVCSILMGSMNAVDNEDPKLYLHLAFTATLFAGVQAAALGLPRLGLIVDFLGHATII 221  
Db 138 AVASLLIGSMLESEEAADDPALYLHVAUTATFFAGVFOALLGVLRGLIVDFLGHATIV 197  
QY 222 GFMGGAATVVCLOQLKSILGLBHFTHGADIISVMRSVFTQTHEWRWESAVLGCVFIFFL 281  
Db 198 GFMGGAATVVCLOQLKGMFLGHFTTATDLVSMSSVFSQTHLWRWESVVMGCGFLFLL 257  
QY 282 STRYFSKKRPRFFWVSAMAPLTSVLGSLLVYTHAEKHGVEVIGELKKGLNPPSLNLV 341  
Db 258 ITRFFSKRPRFFWVSAAAAPLASVITGSLLVYTHAENHGIQVIGLKKGLNPPSATSIN 317  
QY 342 FVSPYMTTAVKTIIVGIIISLAEGIAVGRSPAMYNKYNIDGNKEMIAICTMNVVGSFTSC 401  
Db 318 FSPYMTTAVKTIIVGIIISLAEGIAVGRSPAMYNKYNIDGNKEMIAICTMNVVGSFTSC 377  
QY 402 YLTGTFPSRSVAVNAGCKTAASNIIMSLAVMLTLLFLTPLFHYTPLVVLSAIIIVSAMLG 461  
Db 378 YLTGTFPSRSVAVNAGCKTAMSNVMSVAVMTLLFLTPLFHYTPLVVLSAIIIVSAMLG 437  
QY 462 LIDYEAAILFKVDKDFVVCMSAYIGVVGSGVEIGLVIATVIVSULRVLLFTARPRTFVL 521  
Db 438 LIDYEAAILFKVDKDFVVCMSAYIGVVGSGVEIGLVIATVIVSULRVLLFTARPRTFVL 497  
QY 522 GNIPNSVIYRNMDOYTAARQVGVLRVDSPIYFTNYSYLRERARWIDDEEQCKEG 581  
Db 498 GNIPNSVIYRNMDOYTAARQVGVLRVDSPIYFTNYSYLRERARWIDDEEQCKEG 557  
QY 582 ETSLOQVVIDMSAVGNIDTSGISMLEEVKKITERRELQVLVLPVSEVMKLNKSKFQNH 641  
Db 558 EMGQIVVLDMGAVGSDITSGTSMLEDLAKTLDRLGQIVLANPGSEIMKLDSSKVLAE 617  
QY 642 LGKKWIIYLTVEEAVGACNFNLRSK 666  
Db 618 IGHIEWIFPTVGEAVAECDVFMHSQK 642

RESULT 2  
US-10-424-599-273958  
; Sequence 273958, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Zhou Yihua  
; APPLICANT: Kovalic, David K  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Uses Thereof for Plant Improvement  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 273958  
; LENGTH: 414  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_89404C.1.pap  
US-10-424-599-273958

Query Match 58.0%; Score 2015; DB 15; Length 414;  
Best Local Similarity 94.5%; Pred. No. 9.8e-166;  
Matches 397; Conservative 8; Mismatches 9; Indels 6; Gaps 1;  
QY 261 QTHEWRWESAVLGCVFIFFLSTRYFSKKRPRFFWVSAMAPLTSVLGSLLVYTHAEKH 320  
Db 1 QTHEWRWESAVLGCVFIFFLSTRYFSKKRPRFFWVSAMAPLTSVLGSLLVYTHAEKH 60

QY 321 GVEVIGELKGLNPPSLTNLVFVSPYMTTAVKTIIVGIIISLAEGIAVGRSPAMYNKYN 380  
Db 61 GVQ-----KGLNPPSLTNLVFVSPYMTTAVKTIIVGIIISLAEGIAVGRSPAMYNKYN 114  
QY 381 DGNKEMIAIGTMNVVVGSGFTSCYLTGTFPSRSVAVNAGCKTAASNIIMSLAVMLTLLFLT 440  
Db 115 DGNKEMIAIGTMNVVVGSGFTSCYLTGTFPSRSVAVNAGCKTAASNIIMSLAVMLTLLFLT 174  
QY 441 PLFHYTPLVVLSAIIIVSAMLGILIDYEAAILFKVDKDFVVCMSAYIGVVGSGVEIGLVI 500  
Db 175 PLFHYTPLVVLSAIIIVSAMLGILIDYEAAILFKVDKDFVVCMSAYIGVVGSGVEIGLVI 234  
QY 501 ATIVSVLAVLLFIAPRPTFVLGNIPNSVIYRNVHYQNAKHVPGMLILEIDAPIFANAS 560  
Db 235 ATIVSVLAVLLFIAPRPTFVLGNIPNSVIYRNVHYQNAKHVPGMLILEIDAPIFANAS 294  
QY 561 YLRERITRWIDDEEERIKATGETSLQYVVIDMSAVGNIDTSGISMLEEVKKITERRELQL 620  
Db 295 YLRERITRWIDDEEERIKATGETSLQYVVIDMSAVGNIDTSGISMLEEVKKITERRELQL 354  
QY 621 VLVPVSEVMKLNKSKFQNHKGKWIYLTVEEAVGACNFNLRSKTHPKKDETEGWNV 680  
Db 355 VLVPVSEVMKLNKSKFQNHKGKWIYLTVEEAVGACNFNLRSKTHPKKDETEGWNV 414

RESULT 3  
US-10-425-115-235264  
; Sequence 235264, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 235264  
; LENGTH: 695  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_146143C.1.pap  
US-10-425-115-235264

Query Match 53.3%; Score 1853.5; DB 16; Length 695;  
Best Local Similarity 56.3%; Pred. No. 2e-151;  
Matches 346; Conservative 115; Mismatches 153; Indels 1; Gaps 1;  
QY 42 HOVEVPPPPQPPFKSLKYSIKETFFDDPLRQPKPKPKSKKMLGLQFFPIFEWAPKYTF 101  
Db 70 YKVGVPPEKNLLAEISDAVKETFFADDPLRQYKQDQPSKKIWLGLQHPVPLEWSRHSYL 129  
QY 102 QFLKADLIAGITIASIAIPOGISYAKLANLPPIILGLYSFIPPLIYAMGSSRDLAGTV 161  
Db 130 GKFKGDFIAGLTIASLICIPQDYGYSKLANLPAEVLGYSFVPPPLIYAVMGSSRDIAIGV 189  
QY 162 AVCSILMGSMNSNAVDNEDPKLYLHLAFTATLFAGVQAAALGLPRLGLIVDFLGHATII 221  
Db 190 AVVSLGLTGLQNEIDPKTHPLEYRRLAFTATFFAGVTOALLGFFRLGFIIFLGHAAIV 249  
QY 222 GFMGGAATVVCLOQLKSILGLBHFTHGADIISVMRSVFTQTHEWRWESAVLGCVFIFFL 280  
Db 250 GFMGGAATVVCLOQLKGMFLGHFTTATDLVSMSSVFSQTHLWRWESVVMGCGFLFLL 309  
QY 281 STRYFSKKRPRFFWVSAMAPLTSVLGSLLVYTHAEKHGVEVIGELKKGLNPPSLNLV 340  
Db 310 LVAKYIGKRNKKLFWVSAIAPLTSVITSTFFVYITRADKHGVAIVKNIKNGINPPSASLI 369  
QY 341 VVFSVPMYMTTAVKTIIVGIIISLAEGIAVGRSPAMYNKYNIDGNKEMIAICTMNVVGSFTS 400



Db 370 YFTGPLYATGFKIGIVAGMIGLTAETAIAGRTFAALKDXYRIDGNKEMWALGTNNIVSLTS 429  
Qy 401 CYLTGTPSPRSANNAGCKTAASNIIMSLAVMLTLLFLTPHFTPLVLSAIIYSAML 460  
Db 430 CYVATSGFSRSANNAGCKTAVSNVMSIVVWMLTLLIITPLPKYTPNALISIIISAVL 489  
Qy 461 GLIDYAAIHLPKVDKDFDFVVCMSAYIGVVGFSVEIGLVIAIVISLRLVLLFTIARPRTFV 520  
Db 490 GLIDYSAIILWKVDKLDFLACMGAPFGVIFSSVEYGLLIAVAISLAKIILQVTRPRTL 549  
Qy 521 LGNIPNSVIYRNVEHYQNAKHVPGMLILETDAPYIFANASYLRERTRMIDEEBERIKA 580  
Db 550 LGNLPRTIYRNVEQPDATKVPGLIVRVDSAIYFTNSYVKERILRLWLRDEEQDQ 609  
Qy 581 GETSLQYVIDMSAVGNIDTSGISMLSEVKKITERRELQVLVNPVSEVMKLNKSKFQN 640  
Db 610 KLTKTFLIVDLSPVIDIDTSGIHALLEKALEKKRIQLVLNPGPAVIQKRSAKFTD 669  
Qy 641 HLGKKWIYLTVERAV 655  
Db 670 MIGEDNIFLTVDGAV 684

## RESULT 4

US-10-424-599-207945  
; Sequence 207945, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 207945  
; LENGTH: 659  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURES:  
; NAME/KEY: unsure  
; LOCATION: (1)..(659)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_29800C.1.pap  
US-10-424-599-207945

Query Match 53.2%; Score 1850; DB 15; Length 659;  
Best Local Similarity 55.9%; Pred. No. 3.8e-151;  
Matches 359; Conservative 122; Mismatches 153; Indels 8; Gaps 4;

Qy 41 VHOVEPPPPQPFKSLKYSKETFPPDDPLRQFNKNPKASKKFMGLQFPFPIPEWAPKYT 100  
Db 20 VHOVPPPHKSTQKLGRLKETIFFDDPLRQFNKPKASKKFMGLQFPFPIPEWAPKYN 79  
Qy 101 FQPKADLIAGITIASIAIPQGISYAKLANLPILGLIYSSFIPPLIYAMGSSRDLAGVT 160  
Db 80 LKLPKSDLVSLGTIASLIDIPQMSYAKLASLPILGLIYSSFVPLVYAVLGSSKDLAVGP 139  
Qy 161 VAGSLLMGMLNADVDPNEDPKLYLHLAFTATLPAQVFOAALGLFRGLIVDFLSHAT 220  
Db 140 VSIASLVMSMLHQEVSPTTDPILFLQLAFTSTFLFAGLFOALLGIILRGLFIIDFLSKAIL 199  
Qy 221 IGPWGAATVVCLOOLKSIILGLEHFTHGADIISVMSVFTQTHWESAVLGCVFIFL 280  
Db 200 IGFMAAAIIVSLQOLKSLIGITHTFNQMGILIPVMTSVFNHWSNQITLMGICFLVLL 259  
Qy 281 LSTRYSKKRPRPFVWSAMAPLTSVIIGSLLVYFTTHAEKHGVEVIGELKKGLNPPSLTNL 340

Db 260 LLARHVSIKPKLFWVSAGAPLWCVIIISTLLVFAIKAHQHGISAIGKLOOGINPPSNML 319  
Qy 341 VFPSPMTTAVKTGIVVGIISL-AEGIAVGRSFAMFKYNNIDGNKEMIAIGTMVNVGSFT 399  
Db 320 LPHGSHLGVMTKGLITGLISLTXEGIAVGRFAALKKNYKVDGNKEMMAIGFMNVGSFT 379  
Qy 400 SCYLTTGPPSRSAVNNAGCKTAASNIIMSLAVMLTLLFLTPHFTPLVLSAIIYSAM 459  
Db 380 SCVTTTGAFSRSANNAGCKTAVSNVMSVTVMTLLFLMLPLFQYTPNVVLGAIIIVTAV 439  
Qy 460 LGIDYEAHILPKVDKDFDFVVCMSAYIGVVGFSVEIGLVIAIVISLRLVLLFIARPRTF 519  
Db 440 LGIDLPAACTNIKDKDFVVMVMTAFGLVGLFISVOGGLALAVGLTKLILQITRPKT 499  
Qy 520 VLGNTPNSVIYRNVEHYQNAKHVPGMLILETDAPYIFANASYLRERTRMIDEEBERIKA 579  
Db 500 MLGKTPCTDIYRNLOQYKEAVRIPGFLIILSIEAPINFANITYLNERTLWIEEEDNIKE 559  
Qy 580 TGETSQYVIDMSAVGNIDTSGISMLSEVKKITERRELQVLVNPVSEVMKLNKSKFQ 639  
Db 560 --QLSLRFLVLEMSAVSAVDTSGLISFLKELKATLEKKGVELVLNPLAEVIEKLADEA 617  
Qy 640 N-HLGKKWIYLTVERAVGACNENLRASKTNPKKDETEGWNV 680  
Db 618 NDFIRADNLFTVGEAVA----SLSSAMKQSGSTITEGAHTI 655

## RESULT 5

US-10-437-963-198336  
; Sequence 198336, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 198336  
; LENGTH: 666  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURES:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_94005C.1.pap  
US-10-437-963-198336

Query Match 52.1%; Score 1811.5; DB 16; Length 666;  
Best Local Similarity 53.8%; Pred. No. 8.3e-148;  
Matches 340; Conservative 122; Mismatches 167; Indels 3; Gaps 3;

Qy 43 QVEVPPPPQPFKSLKYSKETFPPDDPLRQFNKNPKASKKFMGLQFPFPIPEWAPKYTQ 102  
Db 36 EYVLSGRPRFASKLWSDLAETFFDDPFRGFGALPPARRAWCAVKYFVPALDWVRYGLD 95  
Qy 103 FLKADLIAGITIASIAIPQGISYAKLANLPILGLIYSSFIPPLIYAMGSSRDLAGVTA 162  
Db 96 KGFDPDLAGITIASIAIPQGISYAKLANLPILGLIYSSFVPLVYAVLGSSNNLAGVTA 155  
Qy 163 VGSLLMGMLNADVDPNEDPKLYLHLAFTATLPAQVFOAALGLFRGLIVDFLSHAT 222  
Db 156 AASLLASIIIEVAADENPQLYLQFYTAAPFTGLFQALGVFRGLIVDFLSRSTIG 215  
Qy 223 FNGGAATVVCLOOLKSIILGLEHFTHGADIISVMSVFTQTHWESAVLGCVFIFL 282  
Db 216 FNGGTAMIILOQFKLLGMKHFTTTKTDIISVLHSTHYRHEWKMOSAVLGICFLFLMS 275





```
QY 378 YNIDGNKEMIAIGTMNVGSGFTSCYLTGTPGFRSAVNNYAGCKTAASNIIMSLAVMLTLL 437
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
127 YHIDGNKEMIAIGTMNIFGFTSCYLTGTPGFRSAVNNYAGCKTAASNIIMSLAVMLTLL 186
QY 438 EUTPLPHYTPPLVLSAIIVSAMGLDYEAAIHLFKVDKDFVVCMSAYIGVVFGSVEIG 497
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 FUTPLPHYTPPLVLSAIIVSAMGLDYEAAIHLFKVDKDFVVCMSAYIGVVFGSVEIG 246
QY 498 LVIAIVISLVRLLFTIARPTFVLGNIPNSVIYRNVEHYQNAKHVPGMLILBIDAPIYFA 557
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
247 LVIAIAISLVRLLFTIARP-DXVLGNIPNSVIYRNVEHYPNAXHVPGLMLILBIDAPIYFA 305
QY 558 NASYLREIRTRMIDEBEERIKATGETSLQYVIDMSAVGNIDTSGISMLBEVVKITERR 617
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 NASYLREIRKRWIDEBEERIKA-----GMFDMASAXCMLGEI----- 341
QY 618 LQLVLVNPVSEVMKLNKSKFQNHGLKWIYLTVEEAVGACNPNLRASKTNPKDDETCGW 677
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
342 -----PRSEVMKLNKSKFLDELGQKRYLTVEEAVGACNPNLHSHYKPNPMKDESEGW 394
QY 678 NN 679
Db 395 NN 396

RESULT 10
US-10-437-963-179495
; Sequence 179495, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179495
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(653)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76952C.1.pep
US-10-437-963-179495
```

```
Query Match 49.6%; Score 1723.5; DB 16; Length 653;
Best Local Similarity 53.2%; Pred. No. 3.4e-140;
Matches 328; Conservative 119; Mismatches 165; Indels 5; Gaps 2;

QY 42 HQVEVPPPPFKSLKYSUKETFFDDPLRQFNKPKASKKFMGLQFPFFPIFEWAPKYTF 101
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 YNVGAPPKNLLAEFAGTAKETFFSDPMRRYKQDQPRSKLMLALQHVFPFVEWGRQYTL 87
QY 102 QFLKADLIAGITIASLAIPQGISYAKLANLPILGLYSSFPPIPIIYAMGSSRDLAGVT 161
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 AKFPGDLIAGLTIASLVIPDQIGYAKLANLPPIGLHSSFVPPPIIYALMGTSRELAWGPV 147
QY 162 AVGSLMGSMLSNAVDNEDPKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFLSHATII 221
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 AVISLLGLTLQBEIDSKNPLDYRLATPTATFFAGVTQAALGFCRLGFIIFLSHAALI 207
```

```
QY 222 GFMGAATVVCLOQLKSLGLEHFTHGADIIISVMESVFTQTH---EWRWESAVILGCVFTF 278
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
208 GFMGAATITIALQQLKGLFIANFTKTDIIISVMKSVGNVHHGXQMNWOTILIGASFLA 267
QY 279 FLLSTRYFSKRPFRFWVSAMAPLTSVILGSLIVYFTHAEKHGVEVIGELKKGLNPPSLT 338
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 FLLVA--XGKKKKLFWVPALAPLISVIISTLFVYITRADKQGVAVIKNVKKGINPPSAS 325
QY 339 NLVFSVPMTTAVKTVGIVGIIISLAEGIAVGRSPAMYKNYINIDGNKEMIAIGTMNVGSGF 398
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 LIFFGTPYLLKGFKIGVWAGMISLTEALAVGRTEAGLNDYQIDGNKEMLAGTGMNVVSGM 385
QY 339 TSCYLTGTPGFRSAVNNAGCKTAASNIIMSLAVMLTLELTPPHYTPPLVLSAIIYSA 458
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
386 TSCYIATGCFARSAVNCMAGGKTPMSNIIMSVTVLLALLWITPLPKYTPNATISSIIISA 445
QY 459 MGLGLDYEAAIHLFKVDKDFVVCMSAYIGVVFGSVEIGLVIATIVISLVRLLFIARPT 518
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
446 VLGLDFESAVLIWKVKLDWFACLGAFGLWIFSSVEYGLLIAVVISLIVKLVHVRPT 505
QY 519 FVLGNIPNSVIYRNVEHYQNAKHVPGMLILEIDAPIYFANASYLRERITRWIDEBEERIK 578
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
506 ALLGNLPRTIYRNVEQYPEATKVPGLIIVRVDSAIYFTNSNYVKERMLRLRDEEHOQ 565
QY 579 ATGETSLQYVIDMSAVGNIDTSGISMLBEVVKITERRLEQLVLVNPVSEVMKLNKSKF 638
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
566 EQKLPKIBFLIVDLSPVNDIDTSGIHAFKELLRTUEKQIQILIFANPGAAVIQKURSAKF 625
QY 639 QNHGLGKWIYLTVEEAV 655
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
626 TELIGEKKIULTVGDAV 642

RESULT 11
US-10-425-115-366870
; Sequence 366870, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 366870
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(662)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_97750C.1.pep
US-10-425-115-366870

Query Match 49.6%; Score 1722.5; DB 16; Length 662;
Best Local Similarity 53.9%; Pred. No. 4.3e-140;
Matches 343; Conservative 110; Mismatches 168; Indels 15; Gaps 6;

QY 41 VHQVEVPPPPFPFKSLKYSUKETFFDDPLRQFNKPKASKKFMGLQFPFFPIFEWAPKYT 100
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
35 VHKVAPPARSTARSKMVRVKETFFDDPFRAFQKQPQGTQWLMVAVRYLFPILDWVPSYS 94
QY 101 FOFLKADLIAGITIASLAIPQGISYAKLANLPILGLYSSFPPIPIIYAMGSSRDLAGVT 160
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
95 LSLFKSDLVAGITIASLAIPQGISYAKLASLPPIIGLYSSFFPMPVYAVIYSSRDLAGVP 154
QY 161 VAVGSLMGSMLSNA-VDPNED---PKLYLHLAFTATLFAGVFOAALGLFRLGLIVDFL 215
```

```
Db      155 VSISSLMGKPCXAGREPHRGADAVPAAGLH-----XTLFAGLVQAASLGLILGFIIDFL 210
Qy      216 SHATIIIGFGGAATVCLQOLKILGLEHFTGADIISVMRSVFTOTHEWRWESAVLGCV 275
Db      211 SKATILGFWAGAAIIVALQOLKALIGIVFTTEMGILVPVNASVFHHTSWSQTIILMVC 270
Qy      276 FIFFLISTRYFSKRRPRFFWSAMAPLTSVILGSLAVFYTHAEKKGVEVIGELKKGKLNPP 335
Db      271 FLVFLLSARHVSIRWPKLFWVSACAPLASVTISTLLVFLFKAQNHGISIIGQLKCGLNRP 330
Qy      336 SLTNLVFVSPYMTTAVKTIIVGIIISLAEGIAVGRSFAMKYKNYIDGNKEMIAIGTMNVV 395
Db      331 SWDKLLFDYAILGLTKTGLTGTIIISLREGIAVGRTPASLKQYQIDGNKEMMAIGLMNVV 390
Qy      396 GSFTCYLTGTPPSRAVNVNAGCKTAASNIIMSLAVMLTLLFLTPHYTFLVAVLSII 455
Db      391 GSCSTCYVTGAFPSRAVNVNAGCKTAMSNVIMALTVMVTLFLMPLFVYTTNVVLGAIL 450
Qy      456 VSAMGLIIDEYAAIHLFKVDKDFVVCMSAYIGVWFGSVSEIGIIVIAIVISLVRLLFIAR 515
Db      451 IAAVIGLIDFPAYVHLWKMDKDFLCVCAFAGVIFISVQEGIAVAGSIFRVLWQITR 510
Qy      516 PRFVLGNTPNSVIYRNVHEYQNAKHVPQMLIETDAPYIFANASYLRERITRWIDEBE 575
Db      511 PKMWQGNIKGTDIYRDLHHYKQAQRVSGFLIAIEAPINFANSNYINERIKRWI--EE 568
Qy      576 RIKATGETSLOYVIIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKLNK 635
Db      569 SPEQDKHTELHFLIIDLASVPAIDTSGIAFLDIKKSIEKRGLELVLVNPTGEXHGENTT 628
Qy      636 SK--FQNLHGKGIWYLTVEEAVGACNFNLRSKTNP 669
Db      629 CKRGXXTIIQQIACIYTTGEAI--ASLSALAKWTKP 662
```

## RESULT 12

```
US-10-437-963-174545
; Sequence 174545, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: Thomas J.
; APPLICANT: David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174545
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72475C.1.pap
US-10-437-963-174545
```

```
Query Match      45.0%; Score 1563.5; DB 16; Length 659;
Best Local Similarity 47.7%; Pred. No. 2.6e-126;
Matches 299; Conservative 121; Mismatches 194; Indels 13; Gaps 3;

Qy      50 QPFKSLKYSKLETFPPDPLRQFKNKPASKKFMGLQGFPPFIFEWAPKYTFQFLKADLI 109
Db      29 RPFGAARAALKDTLFPDPPFRGLGGMPPARRAWRVFVPALDWAGVSAASFYIDL 88
Qy      110 AGTTIASLAIPOGISYAKLANLPPIILGLYSSFPPIIYAMGSSRDLAVGTAVGSLLMG 169
```

```
Db      89 AGVTIASISIPOGISYATLAGIPPVIGLYSCFVPLVAVMGSSRNGLGVGPVATSSLLVA 148
Qy      170 SMLSNVADPNEDPKLYLHLAFTATLFAGVFOAALGFRGLGLIVDFLSHATIIIGFMGGAAT 229
Db      149 SIVGKVRASDDORLYTQLVFTSAFTGTVLQAALGLRLGILVDFMSSRPALTIGFMGGTAI 208
Qy      230 VVLCLOOLKSIILGLEHFTGADIISVMRSVFTOTHEWRWESAVLGCVFFFLFLSTRYFSKK 289
Db      209 VIMLQOLKGLGTMHTFTTKTDIVSLRYIFHNTHOMQSTVLGVCFILFLVFTQVRRR 268
Qy      290 RPRFFWVSAMAPLTSVILGSLAVFYTHAEKKGVEVIGELKKGKLNPPSLTNLVFVSPYMTT 349
Db      269 RPKLFWVSAMSPLLVVVVCVFSFLIKGHKGIPVIGLKRGINFSSISQLKQFQEVGV 328
Qy      350 AVKTIIVGIIISLAEGIAVGRSFAMKYKNYIDGNKEMIAIGTMNVVSGTSCYLTGTPPS 409
Db      329 AMKAGFVSGMLAALASGVAVGRSFAMKKERIDGNKEMVAFGLMNLIGSTCYVITTGAFS 388
Qy      410 RSVNVNAGCKTAASNIIMSLAVMLTLLFLTPHYTFLVLSAIVSAMLGLIIDEYAAI 469
Db      389 KTVNVNTHAGCRTAMSNVCMALVVALAPLFRHTPLVALAAIITSSMLGLVKHREIR 448
Qy      470 HLPKVDKDFVVCMSAYIGVWFGSVSEIGIIVIAIVISLVRLLFIARPRTFVLGNI-- 524
Db      449 RLYEVDKADFVCAALIGVVFSTMITGLGVAVALSVLRALJHVARPSTSKLGRVSCGS 508
Qy      525 ----PNSVIYRNVHEYQNAKHVPQMLIETDAPYIFANASYLRERITRWIDEBEERIK 579
Db      509 AGAADDHAFCDVAQYPGNATAPSLVLQVAGSPVCFANAEYLRERIAWVEDEE--KA 565
Qy      580 TGETSLOYVIIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKLNKSKFQ 639
Db      566 VAGEDLLVVLIDIGGVTIADSPGIEMLRREVHGLERKGMKMAVTNPRMAVASKVLVSLGA 625
Qy      640 NHLGKGIWYLTVEEAVGACNFNLRSK 666
Db      626 ELVGESWMFLSNGDALAACRYTLOGSK 652
```

## RESULT 13

```
US-10-424-599-195017
; Sequence 195017, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 195017
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_18127C.1.pap
US-10-424-599-195017
```

```
Query Match      43.2%; Score 1500; DB 15; Length 653;
Best Local Similarity 48.0%; Pred. No. 8.1e-121;
Matches 297; Conservative 126; Mismatches 178; Indels 18; Gaps 7;

Qy      48 PPQPFKSLKYSKLETFPPDPLRQFKNKK-----PASKKFMGLQGFPPFIFEWAPKYT 100
Db      38 PPSF-WRVVADSVSKT-----ISHYKHLSSLIDQPCTLLLSVLQVVFPIAWGRNVT 90
Qy      101 FOFLKADLIAGTITIASLAIPOGISYAKLANLPPIILGLYSSFPPIIYAMGSSRDLAVGT 160
Db      91 ATKFRKDLIAGLTIASLCIPQSIGYATLAHLDPQYGLYTSVVPPLIYAVMGTSRSTAIGP 150
```

```
QY 161 VAVGSLMGSMLSNAVDPNEDPKLYLHFAFTATLTFAGVFOAALGLFRLGLIVDFLSHATI 220
Db 151 VAVVSLSSMMKLVDPATDPVGYTKJLLATLTFAGIFQTSFGLRLGLDFLDFLSHAAN 210
QY 221 IGPMSGAAVVCVLCQOLKSTLGLHFHFGADIIISVMRSVFTQTHE-WRMESAVLGCVFIF 279
Db 211 VGFVAGAAVIGLQQLKGLGGLGTHFTTKTDIVSVMKAVWEAVHNPWPNRNFILGCSFLVF 270
QY 280 LLSTRYFSKKRPRFFWVSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKGINPPSLTN 339
Db 271 ILTTRFLGKRKKKFLWLASISPLSVVLTSLVFLTRADKNGVKLVKHVKGGLNPSLHQ 330
QY 340 LVFVSPYMTAVKTVGLVGIISLAEGIAVGRSFAMYNINIDGNKEMIAIGTMVNVGSP 399
Db 331 LDFNPNYIGEVAKIGLVAVVALLTSTIAVGRSFASIKGYQLDGNKEMSIGLNTIGSFT 390
QY 400 SCVLTGTPFSRSVAVNACKTAASNIIMSLAVMLTLLPLTFHYTPLVLSAIVSAM 459
Db 391 SCVATGSRFRTAVNAACTELVSNVMAITVLISLQFLTLLKYTPTAILASVLSAL 450
QY 460 LGLIDYEAAILHLPKVDKDFVVCMSAYIGVFGSVGEIGLVIAIVISVLRLFIARPTF 519
Db 451 PGLIDLSEAYKIWKVDKIDFLACAGAFGLVFASVEIGLVAVVWISFSKIIISIRPGE 510
QY 520 VLGNTPNSVIYRVEHYQNAKHVPGMLILEI-DAPIYFANASYLRETRTWIDEEB-ERI 577
Db 511 TLGKIPGTDLCDFDVQYPMVAVKVPVGMIIIRVKSALLCFANANFVRERIIKWYTEESED 570
QY 578 KATGETSLOVYIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPVSEVMKKLNKSK 637
Db 571 KNSRSTIQVLIDTSLNVIDTSGITALEELHKSLSGQKQLAIANPRQVHKLKVSAN 630
QY 638 FQNLGKWKIYLTVEAVG 656
Db 631 FVGKIGGR-VFLTVEEAVG 648

RESULT 14
US-10-425-114-69272
; Sequence 69272, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69272
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMST02400016G05_FLI.pap
US-10-425-114-69272

Query Match 42.0%; Score 1460; DB 15; Length 575;
Best Local Similarity 49.3%; Pred. No. 2e-117;
Matches 281; Conservative 120; Mismatches 165; Indels 4; Gaps 4;

QY 90 PFIFEWAPKYTFQFLKADLIAGTIIASLAIPQGISYAKLANIPPLIGLYSSFPPIIYAM 149
Db 2 PFILAWGRNYATKPKRDLAGLTIASLCIPQSIGVATLAHLDPQGYLTSVVPPLIYAV 61
QY 150 MGSSRDLAVGTVAVGSLLMGSMLSNAVDNEDPKLYLHFAFTATLTFAGVFOAALGLFRUG 209
```

```
Db 62 MGTSEIRAI GPVAVVSLSSMMKLVDPATDPVGYTKJLLATLTFAGIFQTSFGLRLGL 121
QY 210 LIVDFLSHATIIGPMGGAATVVCLOOLKSIILGLEHFTGADIISVMRSVFTQTHE-WRWE 268
Db 122 FLVDFLSHAAIVGFGAAGVIGLQQLKGLGTHFTTKTDIVSVMKAVWEAVHNPWPNR 181
QY 269 SAVLGCVFIPELLSTRYFSKKRPRFFWVSAMAPLTSVILGSLVYFTHAEKHGVEVIGEL 328
Db 182 NPLIGCSFLVFLITTRCIGKRRKKFLWLASISPLSVVLTSLVFLITRADKNGVKLVKHV 241
QY 329 KKGLNPPSLTNLIVFVSPYMTAVKTVGIIVGIIISLAEGIAVGRSFAMYNINIDGNKEMIA 388
Db 242 KGGLNPSIHLQDFFNPYIGEVAKIGLVAVVALLTSTIAVGRSFASIKGYQLDGNKEMMS 301
QY 399 IGTNVVGSFTSCYLTGTPFSRSVAVNACKTAASNIIMSLAVMLTLLPLTFHYTPL 448
Db 302 IGLTNIIGSFTSCYVATGTSFRTAVNAACTELVSNVMAITVLISLQFLTLLKYTPT 361
QY 449 VVLSAIVSAMLGLIDYEAAILHLPKVDKDFVVCMSAYIGVFGSVGEIGLVIAIVISVL 508
Db 362 ALLASVILSALPGLIDLSEAYKIWKVDKIDFLACAGAFGLVFASVEIGLVAVVWISFSK 421
QY 509 VLLFIARPTTFVLGNTPNSVIYRVEHYQNAKHVPGMLILEI-DAPIYFANASYLRETR 567
Db 422 IILSIRPGTETLGLKPGTDLCDFDVQYPMVAVKVPVGMIIIRVKSALLCFANANFVRERII 481
QY 568 RWIDEEB-BRIKATGETSLOVYIDMSAVGNIDTSGISMLBEVKKITERRELQVLVNPV 626
Db 482 KWTVEESEDDEKNSRSTIQFLIDTSLNVIDTSGITALEELHKSLSGQKQLAIANPR 541
QY 627 SEVMKKLNKSKFQNLGKWKIYLTVEEAVG 656
Db 542 MQVHKKUKVSNFVGKIRGR-VFLTVEEAVG 570

RESULT 15
US-10-425-115-228851
; Sequence 228851, Application US/10425115
; Publication No. US200404021472A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 228851
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_140304C.1.pap
US-10-425-115-228851

Query Match 41.3%; Score 1436; DB 16; Length 405;
Best Local Similarity 67.3%; Pred. No. 1.5e-115;
Matches 268; Conservative 65; Mismatches 61; Indels 4; Gaps 1;

QY 275 VFIFELLSTRYFSKKRPRFFWVSAMAPLTSVILGSLVYFTHAEKHGVEVIGELKKG 334
Db 3 IICFPFLQ-----SKRPKLFMTISAAAPLTSVVLGSLVYLTTHAENHGEVIGLYLKKGLNP 58
QY 335 PSLTNLIVFVSPYMTAVKTVGIIVGIIISLAEGIAVGRSFAMYNINIDGNKEMIAIGTMV 394
Db 59 PSVTSLOQSPPPMMALKTIITGVIALAEGIAVGRSFAMYNINIDGNKEMIAIGTMV 118
QY 395 VGSFTSCYLTGTPFSRSVAVNACKTAASNIIMSLAVMLTLLPLTFHYTPLVVLVSAI 454
Db 119 LGSFLTSCYLTGTPFSRSVAVNACKTAMSNVMSLAVMVTLLPLTFHYTPLVVLVSAI 178
```

Qy	455	IVSAMLGLIDYEAAILHFKVDKDFVVCMSAYIGVWFGSVEIGLVIATIVISVLRVLLFIA	514
Db	179	IVSAMLGLVDFGAALHLWRVDKDFVCAGAYLGVWFGSVEGLVAVAVSLRLVLRFGA	238
Qy	515	RPRTFVLGNIPNSVIYRNVEHYONAKHVPQMLILEIDAPIYFANASYLRERITRWIDEE	574
Db	239	RPRTTVLGNIPGTWVYRRMDQYAAATVPGVLVLRVDAPVYFANASYLRERISRWIDEE	298
Qy	575	ERIKATGETSLQYVIIDMSAVGNIDTSGISMLEEVKKITERRELQVLVNPVSEVMKKLN	634
Db	299	ERTKSQGEIMGVRYVWLDMGAIGSIDTSGTSMLEDELNKSLLDRRGMQIVLANPGSEIMKKLD	358
Qy	635	KSKFQNLGKKWIYLTVEEAVGACNFENLRASKTNPKKD	672
Db	359	SSKVLQIGHEWVFPPTVGEAVASCDYVLHSHKPGMAKD	396

Search completed: September 12, 2005, 19:14:19  
Job time : 1410 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 12, 2005, 18:42:11 ; Search time 251 Seconds  
(without alignments)  
4432.941 Million cell updates/sec

Title: US-10-762-049-18  
Perfect score: 3475  
Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDDETEGWNV 680

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10762049/runat\_12092005\_120740\_7209/app\_query.fasta\_1.839  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10762049@cgn\_1.1.105 @runat\_12092005\_120740\_7209 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DE TIMEOUT=120 -WARN TIMEOUT=30 -TTHREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3475	100.0	2449	US-09-720-317A-17	Sequence 17, Appl
2	2164	62.3	1981	US-09-720-317A-3	Sequence 3, Appl
3	1901	54.7	2067	US-09-720-317A-15	Sequence 15, Appl
4	1844.5	53.1	2279	US-09-720-317A-1	Sequence 1, Appl
5	1799	51.8	2311	US-09-720-317A-19	Sequence 19, Appl
6	963.5	27.7	2022	US-09-720-317A-21	Sequence 21, Appl
7	731.5	21.1	780	US-09-720-317A-7	Sequence 7, Appl
8	687.5	19.8	1791	US-09-252-991A-7241	Sequence 7241, Ap
9	676.5	19.5	1767	US-09-252-991A-7184	Sequence 7184, Ap
10	675.5	19.4	1240	US-09-720-317A-5	Sequence 5, Appl
11	651	18.7	2441	US-09-785-381-4	Sequence 4, Appl
12	650	18.7	4113	US-09-785-381-2	Sequence 2, Appl

13	643.5	18.5	1773	4	US-09-902-540-7746	Sequence 7746, Ap
14	643.5	18.5	4854	4	US-09-902-540-768	Sequence 768, Ap
15	641.5	18.5	2832	4	US-09-949-016-4516	Sequence 4516, Ap
c 16	625	18.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl
c 17	619	17.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl
18	612	17.6	1818	4	US-09-602-787A-615	Sequence 615, Ap
19	586.5	16.9	2858	4	US-09-949-016-5349	Sequence 5349, Ap
20	586.5	16.9	2882	1	US-08-424-567-1	Sequence 1, Appl
21	586.5	16.9	2882	2	US-08-711-928-1	Sequence 1, Appl
22	586.5	16.9	2882	3	US-09-184-937-1	Sequence 1, Appl
23	578	16.6	2682	4	US-09-614-221A-222	Sequence 222, Ap
24	577	16.6	4927	4	US-09-949-016-5827	Sequence 5827, Ap
c 25	571.5	16.4	31147	4	US-09-596-002-25	Sequence 25, Appl
26	567	16.3	2487	4	US-09-248-796A-6643	Sequence 6643, Ap
27	548.5	15.8	1992	4	US-09-875-811-5	Sequence 5, Appl
28	548	15.8	1971	4	US-09-875-811-9	Sequence 9, Appl
29	548	15.8	2420	4	US-09-875-811-13	Sequence 13, Appl
30	545.5	15.7	2040	4	US-09-875-811-1	Sequence 1, Appl
31	489	14.1	1851	4	US-09-543-681A-249	Sequence 249, Ap
32	485.5	14.0	1818	4	US-09-875-811-7	Sequence 7, Appl
33	485	14.0	1797	4	US-09-875-811-11	Sequence 11, Appl
34	482.5	13.9	1866	4	US-09-875-811-3	Sequence 3, Appl
35	464	13.4	8774	4	US-09-949-016-16258	Sequence 16258, A
36	451	13.0	2513	4	US-09-795-927-6	Sequence 6, Appl
37	451	13.0	3749	4	US-09-795-927-8	Sequence 8, Appl
38	449	12.9	1608	4	US-09-252-991A-5234	Sequence 5234, Ap
39	449	12.9	1911	4	US-09-252-991A-5199	Sequence 5199, Ap
40	426.5	12.3	1419	4	US-09-902-540-6127	Sequence 6127, Ap
41	426.5	12.3	2070	4	US-09-902-540-286	Sequence 286, Ap
c 42	402.5	11.6	1230230	4	US-09-438-185A-1	Sequence 1, Appl
c 43	389.5	11.2	1230025	4	US-09-198-452A-1	Sequence 1, Appl
c 44	372	10.7	1116	4	US-09-252-991A-7460	Sequence 7460, Ap
45	369	10.6	1701	4	US-09-543-681A-372	Sequence 372, Ap

ALIGNMENTS

RESULT 1  
US-09-720-317A-17  
; Sequence 17, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 17  
; LENGTH: 2449  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-720-317A-17

Alignment Scores:  
Pred. No.: 0 Length: 2449  
Score: 3475.00 Matches: 680  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-762-049-18 (1-680) x US-09-720-317A-17 (1-2449)

QY 1 HsGluLeuAlaArgThrLeuSerTyrIleThrHisIleCysLeuLeuArgenThrIle 20  
|||||  
Db 2 CACGAGCTAGCTGGCACATTAAGTTATATACACATATTGCTTGCTTAATACTATT 61

QY 21 IleGluAspMetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArg 40  
DB 62 ATTGAAGATATGGGAGGTAGATTTATGAGTACCCCTTTGGSCATGAACAACCTTTGAGAGA 121  
QY 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60  
DB 122 GTGCACCAAGTCGAGGTTCCACCCGACACCGCTTTTCAAGTCTCTAAAGTACTCTTTG 181  
QY 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80  
DB 182 AAGGAGACTTTCTCCCTGATGACCCCTTTGAGGAGCTCAAGAAACAAGCAGCTTCCAAG 241  
QY 81 LysPheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThr 100  
DB 242 AAGTTCAAGCTTGCGCTTCAGTTCTTCTCCCATTTTCGAATGGGCTCCCAAAATACAC 301  
QY 101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120  
DB 302 TTTCACTTCTTGAAGCTGACCTCATAGCTGGCATCACCATCGCTAGCTTGGCCATTCT 361  
QY 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSer 140  
DB 362 CAGGGCATCAGTTATGCCAAGCTCGCAACCTCCCTCCAATCTTGCACTATATTCGAGC 421  
QY 141 PheIleProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160  
DB 422 TTTATACCAACCATGATGATGATGGTGGTAGCTCGAGGGATTTGGCAGTGGGGACT 481  
QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180  
DB 482 GTGCGGGTTGGATCGCTCTGATGGGTCGATGTTGAGTAATGCCGCTTGATCCCAATGAA 541  
QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200  
DB 542 GACCCAAAGCTTTACCTCCACCTGGCTTTACAGCTACATATTTGCTGGTGTGTTTCAG 601  
QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220  
DB 602 GCTGCGCTTGGGCTGTGTTAGTGGGGTGTATCGTGGATTTCTGTGCATGCAACCATPA 661  
QY 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240  
DB 662 ATAGGTTTCATGGGAGGAGCAGCACGGTGGTGTGTCTGCAGCAACTTAANAATCGATTCT 721  
QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260  
DB 722 GGCCTTGAGCATTTCAACCATGGAGCTGATATCATATCAGTGATGGCTCTGTGTTTCACC 781  
QY 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280  
DB 782 CAAACTCATGAGTGGAGGTGGGAAGGTGCTGTGTTAGGATGTCTTTCATTTTCTTTCCTC 841  
QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300  
DB 842 CTTAGCACAAAGTACTTCAGCAAAAACGACCAAGGTTTTTTGGTGTTCAGCAATGGCG 901  
QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320  
DB 902 CCATTGACGCTCCCTATATTTAGGAAGTCTCTTGGTTTATTTCACTCAACGCGAAGAC 961  
QY 321 GlyValGluValIleGlyLeuLeuLysLysGlyLeuAsnProSerLeuThrAsnLeu 340  
DB 962 GGTGTTGAAGTATAGGAGAACTGAAGAAGGGTTTGAATCCACCATCTCAAAATCTG 1021  
QY 341 ValPheValSerProTyrTrpThrThrAlaValLysThrGlyIleValValGlyIleIle 360  
DB 1022 GTATTTGTGCGCTTACATGACTACAGCTGTCAAAACTGGCATTTGCTGTTGGCATCAT 1081  
QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380  
DB 1082 TCACCTTGCAGGAAGTAAGCAGTAGGAAGCTTTTGCAATGTATAAAAATTACAATATT 1141  
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400

DB 1142 GATGCAACAAGAGAGATGATAGCTATTGGACCATGAACGTAGTTGGTTCTTTCACTCT 1201  
QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420  
DB 1202 TGCTACCTCACAAAGGACCAATTTTCGGGTTCGGCTGTGAACATAAAGCTGGATGCAAG 1261  
QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440  
DB 1262 ACAGCAGCTTCCAACATTAATAATGTCACTTGCAGTAATGTTGACATTTGTTCTCTGACA 1321  
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460  
DB 1322 CCCTTGTTCATTACACTCCCTGGTGGTCTATCAGCTATTTATCGTATCTGCAATGCTT 1381  
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480  
DB 1382 GGACTCATAGATTATGAAGCAGCCATCCATCTATTAAAGTTGACAAATTTGACTTTGTG 1441  
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500  
DB 1442 GTGTCATGAGTGCATACATTTGGCGTGGTCTTTGGCAGTGTGAAAATTTGGCTTAGTCATA 1501  
QY 501 AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal 520  
DB 1502 GCTATTGTAAATATCTGTACTTCGGGTACTTCTATTATTATTCAGGCCCAAGGACATTCGT 1561  
QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540  
DB 1562 TTTGGGCAACATTTCCAAATTTCTGTGATATACCGAAATGTTGAGCACTATCAAAATGCAAAA 1621  
QY 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560  
DB 1622 CATGTTCTGGAAATCTAATTTCTAGAGATTGATGCCAAATTTACTTTGCCAATGCCAGC 1681  
QY 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580  
DB 1682 TATTTAAGAAAGAGATCACAGGTGGATTGATGAAGAAGAAGAAGAAATTAAGCTACA 1741  
QY 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600  
DB 1742 GGGGAGACTAGTTTCAGTATGTTTATAATTGATATGAGTGTCTGTTGGAAACATTGATACA 1801  
QY 601 SerGlyIleSerMetLeuGluValLysLysIleThrGluArgArgGluLeuGlnLeu 620  
DB 1802 AGTGAATAAGTATGCTCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTT 1861  
QY 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640  
DB 1862 GTTTTGGTCAATCCTGTAAAGTGAAGTGAAGAAACTGAACAAATCGAAGTTCCANAAT 1921  
QY 641 HisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPhe 660  
DB 1922 CATTTAGGGAAGAAATGGATCTATCTGACTGTTGAAGAGGGCGTTGGAGCATGCAACTTC 1981  
QY 661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTyrAsnAsnVal 680  
DB 1982 AATCTACGTGCAAGCAAAACGAAACCAAGAAAGATGAACACAGAGGGTTGGAAACAATGTG 2041

## RESULT 2

US-09-720-317A-3  
; Sequence 3, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998

```

; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Zea mays
US-09-720-317A-3

```

**Alignment Scores:**

Pred. No.:	2.17e-258	Length:	1981
Score:	2164.00	Matches:	403
Percent Similarity:	86.0%	Conservative:	86
Best Local Similarity:	70.95%	Mismatches:	79
Query Match:	62.27%	Indels:	0
DB:	4	Gaps:	0

US-10-762-049-18 (1-680) x US-09-720-317A-3 (1-1981)

Qy	105	LysAlaAspLeuIleAlaGlyThrIleAlaSerLeuAlaIleProGlnGlyIleSer	124
Db	8	GAAGTCGACCTCATCGCGGCATACCATGCCAGCCTCGCATCCCGCAGGCGATCAGC	67
Qy	125	TyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIleProPro	144
Db	68	TACGCCAAGCTGGCCAACTCGCGGCCGGTCTCGACTTACTCGAGCTTGTGCGCGCG	127
Qy	145	LeuIleTyrAlaMetMetGlySerSerAspLeuAlaValGlyThrValAlaValGly	164
Db	128	CTGGTGTACGCGCTGATGGGAGCTCCAAGACACTGGCGTGGGACGGTGGCGGTGGCG	187
Qy	165	SerLeuLeuMetGlySerMetLeuSerAsnAlaValAlaAspProAsnGluAaspProLysLeu	184
Db	188	TGCGTGCTCATCAGCTCCATGCTCGCGCAGGAGTGTGCGCGACGGAGAACCCCGTGCTC	247
Qy	185	TyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuGly	204
Db	248	TACCTGCACCTGCCTTCAACGCCACCTCTTGC CGCGGCTCTTCAGAGCCTCGCTCGGC	307
Qy	205	LeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMet	224
Db	308	CTCCTCAGGTTGGGGTTTCATCTGAGACCTGCTGCGCAGCGACGATCGTGGGGTTCA	367
Qy	225	GlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis	244
Db	368	GCGCGCGCGCAGCGTGTGCTGCTGAGAGCTGCTGAAGGCGCTGCTGGCGCTCTG	427
Qy	245	PheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGlu	264
Db	428	TTCAACACCTCCACCGAGCTCGTCTCCGTATCGAATCCGCTTTCAAGCCAGACACCA	487
Qy	265	TrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeuLeuSerThrArg	284
Db	488	TGGCGGTGGGAGAGCGTCTGCTCGGCTGCGGCTTCTCTTCTCTCTCTCTCTCTCT	547
Qy	285	TyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSer	304
Db	548	TTCAACGACAGAGAGCGTCCCAGAGCTGTCTGGATCTCCCGCGCGCGCGCTTGACGT	607
Qy	305	ValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyValGluVal	324
Db	608	GTCTGCTCGCGAGCGTTCGTGTACCTCAACGAGCTGAAACCAACGCGCATCGAAGTG	667
Qy	325	IleGlyLeuLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSer	344
Db	668	ATCGGTTACTCTAAGAAAGGCGTGAATCCACCGTGGTGACAAAGCCTTGCAATTTCT	727
Qy	345	ProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGlu	364
Db	728	CCCTACATGATCTCGCGCTCAAGACTGGGATCATACC CGCGCGTCTATGGCCCTCGCG	787
Qy	365	GlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAaspGlyAsnLys	384
Db	788	GGAAATCGCGTGGGAGAGCTTCGCCATGTTTCAAGAACTACCAATGACGGACAACA	847

### RESULT 3

```

US-09-720-317A-15
; Sequence 15, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding S
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/72
; CURRENT FILING DATE: 2000-12-21

```





```
QY 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLys---SerLysPheGln 639
Db 1923 GTGCTTGTCAATCCAACTGGAGAGTGCATGGAGAAATAACAACTGGCAACGAGGCTGAA 1982
QY 640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValAlaCysAsn 659
Db 1983 AACTATTTAGGCCAGATTGCTTGATCTGACCACTGGCGAAGCAATC-----CCTTCA 2036
QY 660 PheAsnLeuArgAlaSerLysThrAsnProLysLysAsp 672
Db 2037 CTTTCTGCCTTGGCCAAAGATGACAAAACCCCTAAATGGAT 2075
RESULT 5
US-09-720-317A-19
; Sequence 19, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 2311
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-720-317A-19
Alignment Scores:
Pred. No.: 8,046-213 Length: 2311
Score: 1799.00 Matches: 336
Percent Similarity: 72.30% Conservative: 126
Best Local Similarity: 52.58% Mismatches: 172
Query Match: 51.77% Indels: 5
DB: 4 Gaps: 2
US-10-762-049-18 (1-680) x US-09-720-317A-19 (1-2311)
QY 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
Db 187 GTGTACAAAGTGGCTATCCCTCTCGAAGAACTTGGCCACAGAGTTTACAGAAACATTG 246
QY 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
Db 247 AGGGAGACTTCTTCCACGACAAACCGCTGCGTCAGTATAAGGGCCAAATCCGACCGAGG 306
QY 81 LysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThr 100
Db 307 AGGTTTATGATGGGCTGGAGTCTTGTCTTCTATATTGGGTGGGAGGATACAGT 366
QY 101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
Db 367 CTCACAAGTTCAAAGCGCATCTGATTCGGGATTGACCATTGCAAGTCTCTGTATTCTCT 426
QY 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSer 140
Db 427 CAGGACATTTGGCTATTTCGAAGCTTGCTAATCTGGATCCGCAAGTATGGGCTTACTCCAGC 486
QY 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
Db 487 TTCAATTCCTCCATGATCATGCTGCATGGGTAGCTCAAGGGATACGATGGTCCCA 546
QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
Db 547 GTTGCTGTGTTTCTTTTGTAGGTTCATTCTTCAAGCTGAGGTTGACCATGTCAAA 606
QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
```

```
Db 607 AACAGAGGAGNATACATCGCGCTCGCTTTCACGGCAACCTTCTTCGCTGGTATCACTCAA 666
QY 201 AlalAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
Db 667 GCAGCCTTAGGATTTCTAAGGTTAGGATTCCTTATAGAGTTCTTGTGCGATGCTGCGGATT 726
QY 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
Db 727 GTCGATTCATGGGGGGAGCTGCCATTACTATTGCCCTGCAGCAGCTGAATAACGTGTGTG 786
QY 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
Db 787 GGCATCGCAAACTTTACAGGAAACCCGACATAGTTTCTGTCATGGAATCTGCTCGGAGA 846
QY 261 GlnThrHisGlu---TrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePhe 279
Db 847 TCAGTTTCATCAGCGGTGGAACTGGCAGACAATTTGATTGGCGTATCTTCTCTGCTTTTC 906
QY 280 LeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMet 299
Db 907 CTTCTGTTTGGCAAGTACATCGGAAGAAAGAAAGAAAGAGCTTTTCTGGGTGCCAGTAAT 966
QY 300 AlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLys 319
Db 967 GCTCTTAATAATTTCAAGTATCTAGCAACATTTTGTATACATTAATCTGTCGCGGACAAG 1026
QY 320 HisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsn 339
Db 1027 CAAGGAGTTTCAGATAGTGAAGCACATTTGAACAGGGAATCAACCCATCATCATGACACA 1086
QY 340 LeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIle 359
Db 1087 ATTTATTTCACCGCGCCCATTTGTTGCAAAAGGTTTCAAGATCGGTGTTGTTTGGCGGATA 1146
QY 360 IleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsn 379
Db 1147 GTTGCTTTGACAGAACTGTAGTATTTGGAAGGACATTTGCTGTATGAAGGACTACCCAG 1206
QY 380 IleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThr 399
Db 1207 TTAGATGGAACACAGGAGATGTTAGTCACTTGGAAACCATGAACATAGTAGGTCAATGACA 1266
QY 400 SerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCys 419
Db 1267 TCTTGTCTATGTCACAAACAGGTTCCTTTCAGCTTCGGCAGTTAACTTCATGCTGGCTGC 1326
QY 420 LysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeu 439
Db 1327 AAGACTCCTGTATCCAAATGGTTATGTCAGTAGTGGTTCTTCTTACCTTGTGGTCATC 1386
QY 440 ThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMet 459
Db 1387 ACACCGCTATTCAAAATATACACCGAATGCAATCCTTAGGTCGATCATTTATTTCTCGG 1446
QY 460 LeuGlyLeuIleAspTyrGluAlaIleHisLeuPheLysValAspLysPheAspPhe 479
Db 1447 ATCGCGCTTGTGGATCTACGAAGCAGCAATTCATCTGCGAAAGTTGACAAATTTGACCTTC 1506
QY 480 ValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVal 499
Db 1507 ATTGCTTCGATGGGAGCTTTTTCGGTGTGTTTGTATTCGTTGAGATGGCCTCTTGTG 1566
QY 500 IleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPhe 519
Db 1567 ATTTGCTGAGCAATCTCTATTTGCCAAAATACTTCTTCAAGTAAACAGGCGAAGGACAGC 1626
QY 520 ValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAla 539
Db 1627 CTACTTGGAAACCTTCCCGGCACCACTATATACCGGAACATCAGCCAGTATCCAGAGCA 1686
QY 540 LysHisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAla 559
```



```
Db      1687 AAATTTACTCTCGGGTGGTGATGTGAGGGTTGATTCTGCTATTATTTTCCAACTCT 1746
QY      560 SerTyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAla 579
Db      1747 AATTAGCTCGAGAAAGAAATCTTAGTGGCTGACAGACGAAAGACAGAGTAAAGCA 1806
QY      580 ThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAsp 599
Db      1807 GTGGGATTGCTAAATCAGTTTCTCTGATTGTGGAATGTGCGCGTCTACGACATCGAT 1866
QY      600 ThrSerGlyIleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGln 619
Db      1867 ACAAGCGGCATACATCTCTTGAAGATCTATACAAGAACTTTCAGAAAAAGATATGCGAG 1926
QY      620 LeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGln 639
Db      1927 CTCATCTCTGCGAATCTGCTGCTCGTCGTCATAGAAAACTGCAACGCTCGAAGCTCACCC 1986
QY      640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 659
Db      1987 GAGCACATTGGAAGCAGCAATATATCTCTCGCGTCTCTGAGCTGTGCGATTCTGTACG 2046
QY      660 PheAsnLeuArgAlaSerLysThrAsnPro-LysLysAspGluThrGluGlyTyr 677
Db      2047 -----ACGAAGTCGATGCGAGAACCGTGAGCGAAGTAGTTGCGGAGGAATGG 2092

RESULT 6
US-09-720-317A-21
; Sequence 21, Application US/09720317A
; Patent No. 6696292
; GENERAL INFORMATION:
; APPLICANT: Stephen M. Allen
; APPLICANT: Saverio C. Falco
; APPLICANT: Catherine J. Thorpe
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167
; CURRENT APPLICATION NUMBER: US/09/720,317A
; PRIOR FILING DATE: 2000-12-21
; PRIOR FILING DATE: 14-07-1998
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-720-317A-21

Alignment Scores:
Pred. No.:      6,81e-109      Length:      2022
Score:          963.50      Matches:      206
Percent Similarity: 56.30%      Conservative: 129
Best Local Similarity: 34.62%      Mismatches: 243
Query Match:      27.73%      Indels:      17
DB:              4          Gaps:         7

US-10-762-049-18 (1-680) x US-09-720-317A-21 (1-2022)
QY      91 ProIlePheGluThrAlaProLysTyrThrPhe---GlnPheLeuLysAlaAspLeuIle 109
Db      66 CCGTGTTCGGTGGATGCGGAGCTACAGATGGAAGGAGGACTTCCAGGCGGACCTCGCC 125
QY      110 AlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAla 129
Db      126 GCGGGCATCACTGTGCGGCTCATGCTTGTCCTCAGGCAATGTCATATCAAGAGCTGGCT 185
QY      130 AsnLeuProIleLeuGlyLeuTyrSerSerPheIleProProLeuIleTyrAlaMet 149
Db      186 GGGCTTCACCAATTTAGGCTCTACACAGGCTTTGTCCCATATTGCTCAGCGATT 245
QY      150 MetGlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuMetGly 169
Db      246 TTTGGTCTCTACGACAAATTAGCAGTAGTCCAGTGGCACTTGCTCTCTCTGCTAGTGCC 305
```

```
QY      170 SerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeuAla 189
Db      306 AATGTTCTTGGGGTATAGTT-----AATTTCATCTAGTACGTGTACACGGAATTAGCC 359
QY      190 PheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuLeuGlyLeuPheArgLeuGly 209
Db      360 ATATTATTGGCATTCATGTTGGAATACCTGGAATGCTTGAATGCTTGAAGCTTGAAGCTTGC 419
QY      210 LeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThr 229
Db      420 TGGCTTATTTCGTTTTCATTAGCCATTCTGTAATATCTGGAATTCATACAGCTTCGGGCATC 479
QY      230 ValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAla 249
Db      480 GTAATTTGGTTTGTCCCAATCAAGTATTCTTGGGTACAGT---GTTACAGAAGAGTAGC 536
QY      250 AspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrpGluSer 269
Db      537 AAAAATTATACCACTTATTGAGAGTATAAATGCTGGAATAGATCAGTTCTCTGGCCTCCA 596
QY      270 AlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSerLysLys 289
Db      597 TTTGTAATGGGATCAGCGTTTCTTGTATTCTTCTTAATAATGAAGAGCTAGGAAAAACA 656
QY      290 ArgProArgPhePheThrValSerAlaMetAlaProLeuThrSerValIleLeuGlySer 309
Db      657 AATAAAAAATAGCTTTCCTGAGAGCTTCTGTCCTCACTAACAGCTGTGTTCTTGGAAACA 716
QY      310 LeuLeuValTyrPheThrHisAlaGlyLysHisGlyValGluValIleGlyLeuLeuLys 329
Db      717 TTGTTTGTGAAAAATTTTCGCTCCAACT-----GCCATATCAGTGTGAGTGAATACCG 770
QY      330 LysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerPro----- 345
Db      771 CAAGGCTT-----CCCACT-----TTCTCATTCCTCGAGATTGTGA 809
QY      346 TyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGly 365
Db      810 CATCTGATGTCCCTTAATGCCAACTGCAATACTTATCAGTGGTGTGCTATTTGGAGTCT 869
QY      366 IleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGlu 385
Db      870 GTTGGGATTGCTAAAGCGTTAGTCGGAAGAAATGTTATGATGGACTGCAACCAAGAG 929
QY      386 MetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThr 405
Db      930 TTATTGTGGCTTGGCTTATCAAAATATATGCGGTTTCACTTCTCTGCATATCCTGCTACA 989
QY      406 GlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsn 425
Db      990 GGCTCTCTTTCTAGTCTGCTGTGAATCATGAAAGCGGGGCAAGACTGGATTATCAGGA 1049
QY      426 IleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyr 445
Db      1050 ATCAATAATGGGCATAATAAATTTGCACTGCTCTCTTGTGTTTATGACACCACTATTAT 1109
QY      446 ThrProLeuValValLeuSerAlaIleValSerAlaMetLeuGlyLeuIleAspTyr 465
Db      1110 ATACCTCAGTGTGCATTTGGCTGCCAATTGTGATTTCTGTGTCTCAGTGGCTGTGATTA 1169
QY      466 GluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAla 485
Db      1170 GAAAGGCGCACTCTCTGTGGGTATTGTAAGAAGGATTTCTTCTGTGGCGCATGACA 1229
QY      486 TyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSer 505
Db      1230 TTTACTACAACTTAACCTTTTGGCATTCAGATTGTGTCTTGTGGGTGCGGTTTCG 1289
QY      506 ValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIlePro 525
Db      1290 CTGGCAATTTGTGATCCATGAATCTGCAAAATCCGATATAGCTGTTTGTGGCGCTTGCCT 1349
```



Db 73 CCACTCCCGCGTGGGTACCCGACTCGATAGCTGCTCCACTACCGCCGCGCGCTGGTTC 132  
 QY 105 LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer 124  
 Db 133 CGTCCGGATGTCAGAGCGGGCTGTGGTAGCGCAATCCAGATCCCAACCGCATCGCT 192  
 QY 125 TyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIlePro 144  
 Db 193 TAGCCGACATCCCGCGCTTCCCGCCAGTAGGCTCTATGCGTGCATCTCTCCATCGTC 252  
 QY 145 LeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGly 164  
 Db 253 CTGATCTACGCCCTGATCGCGCAGCTCGCGCAGTTGATGGTAGGCCCGCCGCGCGG--- 309  
 QY 165 SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspPro-----AsnGluAspPro 182  
 Db 310 -----ACCGCGGGATGGTCCGCGCGCATCATCCCGCTGGCGGGCGGCGATCCG 360  
 QY 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202  
 Db 361 CAGCGCTGGTGACCTGTCGATGATCGTCGCGATCATGTCGGGTGTTCTCCATCGTC 420  
 QY 203 LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGly 222  
 Db 421 GCGGCGCTGCGCGCGCGCGGTTTCATCGCAGCTTCTCTCGCAGCATCTCTGGTTCGCG 480  
 QY 223 PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeu 242  
 Db 481 TACCTCAAGCGCATCCGCTGAGCTGCTGGTCCGGCAACTGGGCAAGCTGTTCCGCTAC 540  
 QY 243 GluHisPheThrHisGly-----AlaAspIleIleSerValMetArgSerValPheThr 260  
 Db 541 GAGCGCGCAGCAGCGCTTCTGTCGCGCGCATCTCGCGCTGTGGAGAACCTGCTG--- 597  
 QY 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280  
 Db 598 -----CACATCATTTGGCGCAGCTGCTCGGAGCTCTCGCTCTTGTGTGATG 648  
 QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300  
 Db 649 GTGCTGCTG-----CCGCGGGCTTC---CCGAGTGTCCCGG 684  
 QY 301 ProLeuThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGlyLysHis 320  
 Db 685 GCGCTCGCGCGCTCTTCTGCGCAGCTCGCGCGCGCTGTCGGCGCTGGATCGCTAC 744  
 QY 321 GlyValGluValIleGlyLeuLysLysGlyLeuAsnProSerLeu----- 337  
 Db 745 GCGGTGGAACTGCTGGCGAGGTACCGCGCGGCTG-----CCGCAACTGAGCTGGCGG 798  
 QY 338 ---ThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleVal 356  
 Db 799 CAGACGAGCTGAGGAACCTGAAGACCTGCTGGCGAGCC-----ACCGCATCAGC 852  
 QY 357 ValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLys 376  
 Db 853 GTG-----GTCAGTCTTCAGCGCATGCTCACCGCGCGAGCTTCGCGCGCGCTCAC 906  
 QY 377 AsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGly 396  
 Db 907 GGCTACAGCATCAACCCCAACCAAGAAATTCGTGCGCTCGGCTGGCCCAACATCGCGCC 966  
 QY 397 SerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsn 416  
 Db 967 GGGGTCTCCAGGGCTTTCGCATCAGCGCGCGCGACTCAGCAGCCGCGGTGAACGACATG 1026  
 QY 417 AlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeu 436  
 Db 1027 GTCCGGCGGACAGACCACTGTCGCGGTGGTCCGCGCGCTGGTATCGCCGCCACCTG 1086  
 QY 437 LeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleVal 456  
 Db 1087 TTGCTGCTGAACAGCCCTGGGCTGGGTGCGGATGTCGCGCGCTCGGTGCGGTCTGTTG 1146

QY 457 SerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLys 476  
 Db 1147 CTGCGCGCGTGGGGCTGATCAGCTGCAGCGCTGAAGGGCTTCTGGAAGCTCAGCCGC 1206  
 QY 477 PheAspPheValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIle 496  
 Db 1207 TTTCAGTTTCAAGCTTCTGCTGCTGACCACTGTCGCGCTGCTCAGGCTCGCGCTGCG 1266  
 QY 497 GlyLeuValIleAlaIleValIleSerValLeuArgValLeuPheIleAlaArgPro 516  
 Db 1267 GGAATCTTCTGCTGCGCTGATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326  
 QY 517 ArgThrPheValLeuGly-----AsnIleProAsnSerValIleTyrArgAsnVal 533  
 Db 1327 AGCAGCGCGTGTCTGCTGCTGATGACGCACTGACGCGCGCATGTC-----GAGCTG 1377  
 QY 534 GluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaPro 553  
 Db 1378 GCCAAGTATCCCGCAGCCACCACTCTTCCCGCGCTGGTATCTACCGCTTCAGCGCCCC 1437  
 QY 554 IleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArgTrpIleAspGluGlu 573  
 Db 1438 CTGCTGTTCTTCAAGCGCGATCTTCAAGCAGCGGGTCTGCGCTGGTGGAC----- 1491  
 QY 574 GluGluArgIleLysAlaThrGlyLeuThrSerLeuGlnTyrValIleIleAspMetSer 593  
 Db 1492 -----GGCAGCGCGCGCGCAATGCGTGTCTGCTCAACCGCGAA 1530  
 QY 594 AlaValGlyAsnIleAspThrSerGlyLeuSerMetLeuGluGluValLysIleThr 613  
 Db 1531 GCCATGACCAACTCAGCATCAGCGCGCTGGCGACCTCCAGAGGTGCAACAGATCTCTC 1590  
 QY 614 GluArgArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLysLysLeu 633  
 Db 1591 AAGCGCCAGGGGTGACCTGCTGCGCGGGTGACCGCGCGCTGCTGCTGCTGCTGCTGCTG 1650  
 QY 634 AsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGlu 653  
 Db 1651 CAACGCTCGAGCATCTCGCGGAGATCAAGCGCGCTGGTGTTCAGCTCGGTGCGCTCC 1710  
 QY 654 AlaValGlyAlaCysAsnPheAsnLeuArg 663  
 Db 1711 GGGGTCAAGTCTATCGCTACTGCTGCTGCGC 1740

## RESULT 9

US-09-252-991A-7184  
 ; Sequence 7184, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; ORGANISM: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 7184  
 ; LENGTH: 1767  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-7184

Alignment Scores:  
 Pred. No.: 2,91e-73 Length: 1767  
 Score: 676.50 Matches: 172  
 Percent Similarity: 50.1% Conservative: 117  
 Best Local Similarity: 29.86% Mismatches: 242  
 Query Match: 19.47% Indels: 45

DB: 4 Gaps: 13  
US-10-762-049-18 (1-680) x US-09-252-991A-7184 (1-1767)  
QY 99 TyrThrPheGlnPheLeuLysAlaAepLeuIleAlaGlyIleThrIleAlaSerLeuAla 118  
DB 11 TACCGCGCGCTGGTTCGTCGATGTCAGGCGCGGCTGTCGTAGCGCATCCAG 70  
QY 119 ileProGlnGlyIleSerTyrAlaLysLeuAlaAenLeuProProIleLeuGlyLeuTyr 138  
DB 71 ATCCCCACCGCATCGCTACGCCAGATCGCGGCTTCCGCCCCAGGTAGGCTCTAT 130  
QY 139 SerSerPheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaVal 158  
DB 131 CGGTGCATCCTCGCGATGTGATACGCCCTGTACGGCAGTCGCGGAGTTGATGGTA 190  
QY 159 GlyThrValAlaValGlySerLeuMetGlySerMetLeuSerAsnAlaValAspPro 178  
DB 191 GCGCGCGACGCGCG-----ACCGCGCGATGTCGCGCGCCATCACCCCG 238  
QY 179 -----AsnGluAspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAla 196  
DB 239 CTGGCGCGCGGCGATCCGCGCGCTGTCGTCGACCTGTCGATGATCGTCGATCATGGTC 298  
QY 197 GlyValPheGlnAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSer 216  
DB 299 GGGCTGTTCTCATCGTCGCGCGCTGCGCGCGCGGCTTCATCGCCAGCTTCCTCTCG 358  
QY 217 HisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnLeu 236  
DB 359 CGACCGATCCTGGTGGCTACTCAACGCGATCGCGCTGAGCTGTCGTGGGCACTG 418  
QY 237 LysSerIleLeuGlyLeuGluHisPheThrHisGly-----AlaAspIleIleSerVal 254  
DB 419 GCGAAGCTGTCGGCTACGAGCGCGACGAGCGGCTTCGTCGCGGATCTCTCGCGCTG 478  
QY 255 MetArgSerValPheThrGlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCys 274  
DB 479 CTGGAGAACCTGCTG-----CACATCCATTGGCCGACGCTGATCTCGCGACG 526  
QY 275 ValPheIlePhePheLeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePhe 294  
DB 527 CTCTCGCTCTGCTGATGTCGTGCTG-----CCGCGCGCGCTC 565  
QY 295 TrpValSerAlaMetAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPhe 314  
DB 566 ---CCGCGATTGCGCGCGCTCTCGCGGCTTCTCTGCGCAGCTCGCGCGCGCTG 622  
QY 315 ThrHisAlaGluLysHisGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnPro 334  
DB 623 CTGGCGCTGGATCGCTACGCGCTGGGAACCTGCTTGGCGAGGTACCGCGCGGCTG----- 676  
QY 335 ProSerLeu-----ThrAsnLeuValPheValSerProTyrMetThrAla 350  
DB 677 CGCAACTGAGTGGCGCGCAGACCGCTGGAGAACTGAAAGCTGTCGCGCGACGCC 736  
QY 351 ValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArg 370  
DB 737 -----ACCGCATCACGGTG-----GTCAGCTTCTGCGAGCGCATGTCACCGCGCGC 784  
QY 371 SerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGly 390  
DB 785 AGCTTCGCGCGCGCTCACGCTACAGCATCAACCCCAACCAAGAAATTCGTCGCCCTCGGC 844  
QY 391 ThrMetAsnValValGlySerPheThrCysTyrLeuThrThrGlyProPheSerArg 410  
DB 845 CTGGCCCAATCGCGCGCGGCTCTCCAGGGCTTCGCCATCAGCGCGCGCCGACTCACGC 904  
QY 411 SerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeu 430  
DB 905 ACCGCGGTGAACGACATGTCGCGCGCAAGACCAACTGTCGCGCTGTCGCGCGCTG 964  
QY 431 AlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValVal 450

DB 965 GTGATCGCCGCCACCTGTTGCTGTAAGCCCTGGGCTGGGCTCCCATGCGCGCG 1024  
QY 451 LeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAlaIleHis 470  
DB 1025 CTCGTGCGGCTGTTGCTGCGCGCTGGGCTGATCGACGTCGAGCGCTGAAGGCGC 1084  
QY 471 LeuPheLysValAspLysPheAspValValCysMetSerAlaTyrIleGlyValVal 490  
DB 1085 TTCTGGAAGCTACGCCGCTTCAGTTCAGCTTCGCTGCTGCTGACACCGCTGGGCTGCTC 1144  
QY 491 PheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeu 510  
DB 1145 AGCGTCGCGTGTCCCGGAAATCTTCGTCGCGCTGACATCGCGTACTGCGCTGCTC 1204  
QY 511 LeuPheIleAlaArgProArgThrPheValLeuGly-----AsnIleProAsnSer 527  
DB 1205 TACTACACTATCGCGCGAGCGCGCTGCTGCGTGGATGCAAGCATCGACGCCAG 1264  
QY 528 VallIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIle 547  
DB 1265 GTC-----GAGCTGCGCAAGTACCGCGAGGCCACCTCTTCCGCGCTGCTGATC 1315  
QY 548 LeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThr 567  
DB 1316 TACCGCTTCGAGCGCGCTGCTGCTTTCACCGCGAGTACTTCAAGCAGCGGCTGCTT 1375  
QY 568 ArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyr 587  
DB 1376 GCCGTGGTGGAC-----GGCAGCGAGCGCGCAATGCG 1408  
QY 588 VallIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 607  
DB 1409 GTGCTGCTCAACGCGCAAGCCATGACCACTCGACATCAGCGCGCTCGCCACCCTCCAC 1468  
QY 608 GluValLysLysIleThrArgArgGluLeuGlnLeuValLeuValAsnProValSer 627  
DB 1469 GAGTGTCAACAGATCCTCAAGCCCGAGCGGTGCACCTGTCGCTGGCGGGTGACCGGG 1528  
QY 628 GluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIle 647  
DB 1529 CAGACGCTGGACCTGCTCAACGCTCGAGCATGCTCGCGCAGATCAAGCGCGCGCTGCTG 1588  
QY 648 TyrLeuThrValGluAlaValGlyAlaCysAsnPheAsnLeuArg 663  
DB 1589 TTCAGCTCGGTGCTCGCGGTGAGTGCCTATCGCTACTGCTGCTGCGC 1636

RESULT 10  
US-09-720-317A-5  
; Sequence 5, Application US/09720317A  
; Patent No. 6696292  
; GENERAL INFORMATION:  
; APPLICANT: Stephen M. Allen  
; APPLICANT: Saverio C. Falco  
; APPLICANT: Catherine J. Thorpe  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167  
; CURRENT APPLICATION NUMBER: US/09/720,317A  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/092,833  
; PRIOR FILING DATE: 14-07-1998  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 1240  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-720-317A-5  
Alignment Scores:  
Pred. No.: 2,05e-73 Length: 1240  
Score: 675.50 Matches: 132  
Percent Similarity: 67.22% Conservative: 69

```
Best Local Similarity: 44.15% Mismatches: 97
Query Match: 19.44% Indels: 1
DB: 4 Gaps: 1

US-10-762-049-18 (1-680) x US-09-720-317A-5 (1-1240)

QY 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
DB 6 AGCCTCAGGAAGCTATCGCGTGGCCGATCTTCGCTCCGTAAGAGGGGTACAGACTC 65
QY 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
DB 66 GACGGCAACAAGAGAGATGCTGGCGATCTCCCAACGTTGCTGCTCTCTGCTCTCG 125
QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
DB 126 TGTATGTGGCAACAGGTTGCTTCTCCGACCGCAGTGAACCTTCAGCGCGGGGCCAGG 185
QY 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuPheLeuThr 440
DB 186 TCGACCGTTTCAAAACATCGTCATCACCGTGTTCGTCCACCTGGAGCTGTTTCATG 245
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
DB 246 AAGCTCTCTACTACACGCGCATGGCGGTGCTCGCTCCATCATCTCTGCGGCTCTCCG 305
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
DB 306 GGAATGATGCATCAAGGAGGCTCGACGATATGAAGATCGCAAGATGGATTTCCTC 365
QY 481 ValCysMetSerAlaTyrIleGlyValPheGlySerValGluIleGlyLeuValIle 500
DB 366 ACCTGCTCGGTGGTGTGGCGTCTGTTGGGTGCGTGGAGATTGGCTTCGAGTT 425
QY 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
DB 426 GCATCTGGCATTTCTTCGCAAGATCATCATACAGTCGCTCGGCTCAGGTGGAGATC 485
QY 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
DB 486 CTTGGCAGGCTACAAGGACAGATATCTTCGACGCGTCAGGAGTACCCCTGAGCCTGC 545
QY 541 HisValProGlyMetLeuIleLeuGluIleAspAlaPro---IleTyrPheAlaAsnAla 559
DB 546 CTAACCTCCGACTGTACTGCTATACGCGTCGACACATCTCTCTCTCTTCATCAAGCC 605
QY 560 SerTyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAla 579
DB 606 ACTTCGCTCAAGAAAGGATCACAGAGTGGGTTTGGGAAGGAGTGGAGACCTCAAAATGGA 665
QY 580 ThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAsp 599
DB 666 AAGCGAGGAGGAGGATACAGCAGTGTCTTGATATGTCAAGTGTGTAAACATCGAC 725
QY 600 ThrSerGlyIleSerMetLeuGluGluValLysIleThrGluArgGluLeuGln 619
DB 726 ACTTCAGGACTCACTGCCTGGAAGAAATACACAGGAGTGTGTGTCTCTTGGCTTACAG 785
QY 620 LeuValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGln 639
DB 786 ATGGCTATAGCCAGTCCGGGATGGAAGGCGATTTGAGAAAGTTCAGAAAGTGTACAGGTGGT 845
QY 640 AsnHisLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCys 658
DB 846 GACAGGGTAGGACAGGACTGATCTTCATGACAGTAGGTGAGCGGTGGAGCGCTGT 902
```

RESULT 11

US-09-785-381-4  
; Sequence 4, Application US/09785381  
; Patent No. 6602992  
; GENERAL INFORMATION:  
; APPLICANT: DALLOS, Peter  
; APPLICANT: ZHENG, Jing

```
; APPLICANT: MADISON, Laird
; TITLE OF INVENTION: A MAMMALIAN PRESTIN
; FILE REFERENCE: 0290-37U1
; CURRENT APPLICATION NUMBER: US/09/785,381
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,461
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2441
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-785-381-4

Alignment Scores:
Pred. No.: 7.8e-70 Length: 2441
Score: 651.00 Matches: 191
Percent Similarity: 45.95% Conservative: 144
Best Local Similarity: 26.20% Mismatches: 270
Query Match: 18.73% Indels: 124
DB: 4 Gaps: 21
```

```
US-10-762-049-18 (1-680) x US-09-785-381-4 (1-2441)

QY 35 MetAsnAsnPheGluArgValHisGlnValGluValProPro----- 48
DB 198 CTGCAGACCATGGATCATCTCTGAAGAAATGAATCCCTGCAGAGACCCAGAGGTACTAC 257
QY 49 ---ProGlnProPhePhe-----LysSerLeuLysTyrSerLeuLys 61
DB 258 GTGGAAGAGGCCCATCTTCACTCATCTCTCTCAAGAGAGCGCTGCACGTCGAAGGACAAA 317
QY 62 GluThrPhePheProAspProLeuArgGln---PheLysAsnLysProAlaSerLys 80
DB 318 GTCACAGAGTCCATTGGAGATGAAGTGAAGAGGAGCATTCACGTGACTCCT-----AAA 371
QY 81 LysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluThrAlaProLysTyrThr 100
DB 372 AAAATAAGAAACATCATTTACATGTTCTGCTCTATCACTAAGTGGTGCACGACATATAA 431
QY 101 Phe---GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIle 119
DB 432 TTCAAGGAGTATGTGTAGGTGACTTGGTCTCGGGCATAAAGCACATGGGGTACTCCAGCTT 491
QY 120 ProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSer 139
DB 492 CCCCAAGGCTTAGCCTTCGCCATGCTGCGACGCGTCCGCTCCGCTGTGTACTCA 551
QY 140 SerPheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGly 159
DB 552 TCGTTTACCCCGTTATCATGTACTGTTCTTTGGAACTCAAGACACATATCTATAGT 611
QY 160 ThrValAlaValGlySerLeuLeuMetGlySerMetLeu----- 172
DB 612 CCTTTTGTCTTATTAGTTGATGATGAGGTGTGGCGTGTGGCGTGTGGCGTGTGGCGTGT 671
QY 173 -----SerAsnAlaValAspProAsnGlu---AspProLysLeuTyr 185
DB 672 ATTGTATCCCGAGGAGGATAAATCAACCAACGGGACAGCCAGAGATGCACCTAAGA 731
QY 186 LeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeu 205
DB 732 GTGAAGTGGCATGTCTGTGTACCTTACTTTCAGGAATCATTCAGTTTGGCTTAGGTGTC 791
QY 206 PheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGly 225
DB 792 TGTAGTTGGATTGTGGCCATATACCTCACGGAGCCATTGGTGGAGGCTTTTACCCT 851
QY 226 GlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuHisPhe 245
DB 852 GCGGCTGTGTCCAGCTGTTCATGCTCCATGTTTAAATACCTGTTTGGGGTCAAAACAAAG 911
```

Qy	246	ThrHisGlyAlaAspIleIleSerValMetArgSer-----ValPheThrGlnThr	262
Db	912	CGGTAC---AGTGGAAATCTTTTTCAGTGGTGTATAGTACAGTTGCTGTGTGCAGAAATGTT	968
Qy	263	HisGluTrpArgTrpGluSerAlaValIleGluCysValPheIlePhePheLeuLeuSer	282
Db	969	AAAAACCTCAACGGTGTCTCCCTAGGCGTGGCCCTGATGGTTTGTGGTTGCTGTGTTGGGT	1028
Qy	283	ThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeu	302
Db	1029	GGCAAGAAATTTAATGAGAGA-----TTTAAAGAGAAATGGCCAGCACCCCAT	1076
Qy	303	Thr-----SerValIleLeuGlySer---LeuLeuValTyrPheThrHisAla	317
Db	1077	CCTCTAGAGTTCTTTGCTGTGGTGAAGGAGCTGGCATTTCTGCAGGATTTAACTACAT	1136
Qy	318	GluLysHisGlyValGluValIleGlyGluLeuLysGlyGlyLeuAsnProPro-----	335
Db	1137	GAGTCTACAGTGTGGATGTGTTGGAACTTCTCTCGGGGTACTTCTCTCGGGCCAAC	1196
Qy	336	-----SerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValLys	352
Db	1197	CCAGACACCAAGCTTGTCCACTGGTGTATGTGAC-----	1232
Qy	353	ThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGlyArgSerPhe	372
Db	1233	--GCCATTGCCATCGGCATCGTTGGATTTTTCAGTAGCATCTCCATGGCCAAAACCTTG	1289
Qy	373	AlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyThrMet	392
Db	1290	GCAAAATGAAGCATGGCTACCAAGTTGTATGGCAATCAGGAGCTCATGTGCTTGGGGATATGC	1349
Qy	393	AsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSerArgSerAla	412
Db	1350	AATCTCATGTGATCTCTCTCCAAACCTTCTCGATTTCTGCTCTCTCGAGGCTT	1409
Qy	413	ValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSerLeuAlaVal	432
Db	1410	GTTCAGGAAGGAACCTGGAGGGAAAACACAGCTTCGAGGTGTTGTTGGCTCGTTGATGATT	1469
Qy	433	MetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeuValValLeuSer	452
Db	1470	CTGTTGTGTCATATTAGCCACCGGATTCCTCTTTGAGTCGTTTACCCAGGCTGCTCTTTC	1529
Qy	453	AlaIleIleValSerAlaMetLeuGlyLeu---IleAspTyrGluAlaAlaIleHisLeu	471
Db	1530	GCCATTGTGATCGTCAACCTGAAAGGAATGTTTCATGCAGTGTCTCAGACCTGCCTTTTTT	1589
Qy	472	PheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGlyValValPhe	491
Db	1590	TGGAGAACCAAGCAAAATAGAGCTGACCATCTGGCTGACCACTTTGTGCTCTCCCTGCTTC	1649
Qy	492	GlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArgValLeuLeu	511
Db	1650	CTCGGCTGGACTACGGACTGATTACCGCGCTGATCATGTCTGTCTGCACAGTGAATTAT	1709
Qy	512	PheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArg	531
Db	1710	AGAACACAGAGTCCAAGCTCAAAAGTCTCTGGGGCAGCTCCCTGCACGGATGTGTACATT	1769
Qy	532	AsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAsp	551
Db	1770	GACATAGATGCATATAGGAGGTGAAGAAATTCCTGGAATAAAAAATATTCCAAATAAAT	1829
Qy	552	AlaProIleTyrPheAlaAsnAla-----	559
Db	1830	GCCCAATTTACTATGCAAAATAGCGACTTGTATAGCAGCGCTTTAAAAAGAAAGACTGGA	1889
Qy	560	-----	564
Db	1890	GTAACCCAGCACTCATATTATGGGAGCGGAAGAAAGCCATGAGGAAGTACGCCCAAGAA	1949
Qy	565	-----ArgIleThrArgTrpIleAsp---	571

D	b	1950	GTTGGAAATGCCAACGTGCGCAATGCTACTGTTGTCANAGTGGATGCAGAAGTAGACCGA	2008
Q	y	572	-----GluiGluGluLuarG-----Ile	577
D	b	2010	GAAAACTCTACAAAACCTGAAGAAGAGGATGATGAAGTCAAATTCCCCCAATAGTCATC	2069
Q	y	578	LysAlaThrGlyGluThrSerLeuGln-----Tyr	587
D	b	2070	AAAAACAACATTTCTCTGAAGAGAGCTGCAGAGATTTTTGCCCCAGGGGAAAAATGTCCACACT	2129
Q	y	588	ValIleIleAspMetSerAlaValGLYAsnIleAspThrSerGlyIleSerMetLeuGlu	607
D	b	2130	GTCATTTCTAGACTTTTCCGAGGTCAAATTTTGTAATCTCTGTTGGAGTGAATACTGTGCC	2189
Q	y	608	GluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProValSer	627
D	b	2190	GGGATTCGAAAGAATAATGAGATGTTTGGAAATTTATGTATATTTAGCAGGATGCAGCCCA	2249
Q	y	628	GluValMetLysLysLeuAsnLysSerLys---PheGlnAsnHisLeuGlyLysLysTrp	646
D	b	2250	CNAGTTGTGTAATGACCTCACCCGCAACAACACTTTTTTGANAATCTGCTTTGAAAAGAGCTT	2309
Q	y	647	IleTyrLeuThrValGluGluAlaVal	655
D	b	2310	CTGTTCCACAGTATCCAGATGCAGTC	2336
<b>RESULT 12</b>				
US-09-785-381-2				
; Sequence 2, Application US/09785381				
; Patent No. 6602992				
; GENERAL INFORMATION:				
; APPLICANT: DALLOS, Peter				
; APPLICANT: ZHENG, Jing				
; APPLICANT: MADISON, Laird				
; TITLE OF INVENTION: A MAMMALIAN PRESTIN				
; FILE REFERENCE: 0290-37UI				
; CURRENT APPLICATION NUMBER: US/09/785,381				
; CURRENT FILING DATE: 2001-02-16				
; PRIOR APPLICATION NUMBER: US 60/183,461				
; PRIOR FILING DATE: 2000-02-18				
; NUMBER OF SEQ ID NOS: 16				
; SOFTWARE: PatentIn version 3.0				
; SEQ ID NO 2				
; LENGTH: 4113				
; TYPE: DNA				
; ORGANISM: Meriones unguiculatus				
US-09-785-381-2				
<b>Alignment Scores:</b>				
Prod. No.: 2,66e-69 Length: 4113				
Score: 650.00 Matches: 189				
Percent Similarity: 46.63% Conservative: 143				
Best Local Similarity: 26.54% Mismatches: 254				
Query Match: 18.71% Indels: 126				
DB: 4 Gaps: 21				
US-10-762-049-18 (1-680) x US-09-785-381-2 (1-4113)				
Q	y	39	GluArgValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyr	58
D	b	311	GAGAGGCTGCAT-----GTCAGAGGAC	331
Q	y	59	SerLeuLysGluThrPhePheProAspAspProLeuArgGln---PheLysAsnLysPro	77
D	b	332	AAAGTCTCAGAGTCCATV-----GGGGATAAGCTGAAGCAGCGGTTCACATGCACCTCCC	385
Q	y	78	AlaSerLysLysPheMetLeuGlyLeuGlnPhePhePheProfilePheGluTrpAlaPro	97
D	b	386	-----AAAAGATAAGAAACATCATTTACATGTTCTTGCCCATCATCTAAGTGGTTGCCA	439
Q	y	98	LysTyrThrPhe---GlnPheLeuAlaAspLeulleAlaGlyIleThrIleAlaSer	116



Db 440 GCTTACAGTTCAGGAGTATGTGTTGGGTGACTTGTGTTTTCAGGCATACAGCACCCGGCGTG 499  
QY 117 LeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProPheLeuGly 136  
Db 500 CTTACAGTTCACCAAGCTTAGCTTGGCTGCGAATGCTTGGCGTGTCTCTCGGTGTTCCGGC 559  
QY 137 LeuTyrSerSerPheLeuProProLeuLeuTyrAlaMetMetGlySerSerArgAspLeu 156  
Db 560 CTGACTCTTCATTTATCTGTTATCATGTACTGTTCTTTGGGACCTCCAGACACATA 619  
QY 157 AlaValGlyThrValAlaValGlySerLeuLeuMetGlySerMetLeu 172  
Db 620 TCTATAGTTCCTTTCGCGCTTATAGTTCATCGTGTGTTGCTGCTGCGCTGGTC 679  
QY 173 -----SerAsnAlaValAspProAsnGlu---AspPro 182  
Db 680 CCGATGACATCGTCATCCGGGAGGAGTGAACGCAACCGCAGCGAGGCGCCGAGAC 739  
QY 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLa 202  
Db 740 CGCGTGAGAGTGAAGAGTGCATGCTGTCCACCTGCTCTCAGGAATCATTCAGTTTGC 799  
QY 203 LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGly 222  
Db 800 CTAGGTGTGTCAGGTTTGGATTGTGCGCATATACCTCACGGAGCGCTGTGTCGAGGG 859  
QY 223 PheMetGlyGlyAlaAlaValValCysLeuGlnLeuLysSerIleLeuGlyLeu 242  
Db 860 TTCACACCGCCCGCGTGCACGCTTCACATCCATGTTGAATACCTGTTTGGGGTT 919  
QY 243 GluHisPheThrHisGlyAlaAspIleIleSerValMetArgSer-----ValPhe 259  
Db 920 AAGACAAAGCGGTAC---AGTGGATCTTTTCGTGTGTATATAGTACAGTGTGTTG 976  
QY 260 ThrGlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePhe 279  
Db 977 CAGAATGTTAAACACCTCAACGTGTGTTCTCAGCGGTGCGCTGTGTTTGTGTTTG 1036  
QY 280 LeuLeuSerThrArgTyrPheSerLysLeuArgProArgPhePheTrpValSerAlaMet 299  
Db 1037 CTGTGTGGTGGCAAGAGGTTTAATGAGAGA-----TTTAAAGAGAAATTGCCA 1084  
QY 300 AlaProLeuThr-----SerValIleLeuGlySer---LeuLeuValTyrPhe 314  
Db 1085 GCACCATTCCTCTAGAGTCTTTTGTGTGTGTGAGTGGAACTGGCATTTCCGGGGGTTT 1144  
QY 315 ThrHisAlaGluLysHisGlyValGluValIleGlyLeuLysLysGlyLeuAsnPro 334  
Db 1145 AACTTGCACAGTTCCTACAGTGTGGATGTGTTGGAACTCTTCTCTGGGGCTACTCCCT 1204  
QY 335 Pro-----SerLeuThrAsnLeuValPheValSerProTyrMetThrThr 349  
Db 1205 CTTGCCAACCCGGACACCGCTTCCTCCACTCGTGTATGTGGAT----- 1249  
QY 350 AlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaValGly 369  
Db 1250 -----CCCATGGCATAGCATGTTGGATTTCAGTGCACAAATTCATGGCC 1297  
QY 370 ArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIle 389  
Db 1298 AAAACCTTGGCGAATAAGCATGCTACCGAGTTGATGGCAATCAGGAGCTCATCGCTTTC 1357  
QY 390 GlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPheSer 409  
Db 1358 GGGATATGCAACTCCATCGGATCTCTCTCCAGACCTTCTCCATTTCTGCTCTTGTCT 1417  
QY 410 ArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMetSer 429  
Db 1418 CGAGGCTGTTTCAGGAGGGAACCTGGAGGGAACACAGCTCGCAGGTTGCTTGGCGCTG 1477  
QY 430 LeuAlaValMetLeuThrLeuPheLeuThrProLeuPheHisTyrThrProLeuVal 449  
Db 1478 CTGATGATTCTGCTGCTCATTTTAGCCACTGGATTCTCTTTGAGTCAITTCGCCCGAGGCT 1537

QY 450 ValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu---IleAspTyrGluAlaLa 468  
Db 1538 GTGCTCTCGCCATTTGATCTGTAACCTGAAAGGGATGTTTATGTCAGTCTCTCAGATCTG 1597  
QY 469 IleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGly 488  
Db 1598 CCCTCTCTTCTGGAGACCAACAGCAATATAGAGCTGACCATCTGGCTTACCACTTGTGTCC 1657  
QY 489 ValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuAsg 508  
Db 1658 TCCTCTGTTCTCGGCTTGGACTACGGACTGATTTACTGCTGTGATCAITCTCTGCTGACT 1717  
QY 509 ValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerVal 528  
Db 1718 GTGATTTACAGAACCCAGAGTCCGAGCTACAGGTCTCTGGGCGAGCTCTCTGACACCGAT 1777  
QY 529 IleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeu 548  
Db 1778 GTATACATTTGACATAGACGATATGAGGAGGTGAAGAAATTCCTGGAATAAAATATTC 1837  
QY 549 GluIleAspAlaProIleTyrPheAlaAsnAla----- 559  
Db 1838 CAGATAAACGCCCAATTTCATGCAACAGTACTTGTATAGCAACGCCCTTAAAGA 1897  
QY 560 -----SerTyr 561  
Db 1898 AAGACTGTGTGAACCCAGCGCTCATATGGGAGCAAGGAAGGCCCATGAGGAAGTAC 1957  
QY 562 LeuArgGlu-----ArgIleThrArgTrp 569  
Db 1958 GCAAAGGAAGTCGAAACGCCCAACATTCGCAACGCGAGCTGTGTCAAAGTGGATGAGAA 2017  
QY 570 IleAsp-----GluGluGluGluArg----- 576  
Db 2018 GTAGATGGAGAAATGCTACGAAGCCCGAAGAGAGGATGATGAAGTAAATATATCCCCA 2077  
QY 577 -----IleLysAlaThrGlyGluThrSerLeuGln----- 586  
Db 2078 ATAGTATCATCAAAACAACATTTCTGAGAGAGCTGCAGAGATTATGCCCCAGACAGAAAT 2137  
QY 587 -----TyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSer 604  
Db 2138 GTCCACACTATCTATCTAGACTTCACAAGTCAATTTTATCGACTCTCTGTGAGTAAAA 2197  
QY 605 MetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsn 624  
Db 2198 ACCCTGGCTGTGATGGTGAAGAAATACGGAGATGTTGGTATTATTATGTACTAGCAGGA 2257  
QY 625 ProValSerGluValMetLysLysLeuAsnLysSerLys---PheGlnAsnHisLeuGly 643  
Db 2258 TCGACGCCACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2317  
QY 644 LysLysTrpIleTyrLeuThrValGluGluAlaVal 655  
Db 2318 AAAGAGCTTCTGTTCCACAGTATCCATGATGATGATGATGATGATGATGATGATGATGAT 2353

## RESULT 13

US-09-902-540-7746  
; Sequence 7746, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825



```
; SEQ ID NO 7746
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-7746

Alignment Scores:
Pred. No.: 3,76e-69 Length: 1773
Score: 643.50 Matches: 162
Percent Similarity: 50.44% Conservative: 124
Best Local Similarity: 28.57% Mismatches: 256
Query Match: 18.52% Indels: 25
DB: Gaps: 7

US-10-762-049-18 (1-680) x US-09-902-540-7746 (1-1773)

Qy 91 ProIlePheGluTrpAlaProLysTyrThrPheGlnPheLeuLysAlaAspLeuIleAla 110
Db 67 CCGGGCTCGGGAGGACCGGGCTATACGGCGCTGTTCGGGGCGGACCTGCTCTCG 126
Qy 111 GlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeuAlaAsn 130
Db 127 GCGCTCACCATTGGTGCAGTCTCATCCCGAGGGCTTGGCTACGCGAGCTCGTGGGC 186
Qy 131 LeuProIleLeuGlyLeuTyrSerSerPheIleProProLeuIleTyrAlaMetMet 150
Db 187 GTGCGTCCAGCAGCGGGCTCTACGCGGGCGTGTGGGAATGCTGGCCCTATGCGCTGTT 246
Qy 151 GlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMetGlySer 170
Db 247 GGGCCCTCGGGACCTCATCATCGGACCGGAGCGGGTCCGCCATCTCATCCCGCGG 306
Qy 171 MetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeuAlaPhe 190
Db 307 GCGCTCGCG--CCGGTGGCGGGGGCGGGCGCGCTTACGATCGCTGGCGGCA 363
Qy 191 ThrAlaThrIleuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeuGlyLeu 210
Db 364 CTGCTGGCGCTGTGTCTGGCGGTGCTGAGCCTGCTCGGAGGGCTGCTCAAGGTGGCGG 423
Qy 211 IleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrVal 230
Db 424 CTGCGGAGTCTCTGTCCAGCCCATCTCATCGGCTATCATATGCGCGGGCTCATC 483
Qy 231 ValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis-----PheThrHisGly 248
Db 484 ATCATCGGACGCGAGCTCGCGGCTCTTCGGGAGGAGCGCCAGCTCCGACAGCTTCTCG 543
Qy 249 AlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrpGlu 268
Db 544 GGTTCAGGTGTTCAGGTGGCCACCCACCTC--GAACGGAGCCACGTC-----CCG 591
Qy 269 SerAlaValIleGlyCysValPheIlePheLeuLeuSerThrArgTyrPheSerLys 288
Db 592 ACCTCTCTGCGGGCTGGGGCTCATACCGCGCTGCTGCTGCTGGGCACTCTCTGCCCC 651
Qy 289 LysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGly 308
Db 652 AAGGTGCCC-----GGTCCGCTCATCTCTGGTGGTCTCTCACC 687
Qy 309 SerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGluLeu 328
Db 688 ACGGTGGCGGAGGGCTGTTCAGTGGAGCATGGGGGCGATCAAGGTGCTGGGCGCCCTTC 747
Qy 329 LysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThr 348
Db 748 GCGCGGAGCCCCCGCTCCAGGCTCTCCGCTCTCTGCTGCTGAGGAC-----GTGCGG 801
Qy 349 ThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaVal 368
Db 802 TCGCTGCTCCCGCGCTTCAGGCTTGGCGCTGCTCAACTACGCACTGCTGCTGAGGCTGAGG 861
Qy 369 GlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAla 388
Db 862 GGCGGATTCAGCGGACGAGGTTCGGCTACCGGCTGGACAGCACCGAGGATTCCTCGGG 921
Qy 389 IleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPhe 408
Db 922 CAGGCGGAGCCACCTCGCTTACGACAGGGGTTCCTCGGCTGACGCGGAGGAGCAG 981
Qy 409 SerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIleMet 428
Db 982 TCGCGCAGCGGCTCAACGCTCTCCATGGGGGGGGAGCCAGCTGCTGGGGGTGCTCGCC 1041
Qy 429 SerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeu 448
Db 1042 GCAGGGGTGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1101
Qy 449 ValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAla 468
Db 1102 GTGACGCTGGGGCCATCGTCTTCGTGGCGGGGTACCTGCTGGAGTTCGAGCCATC 1161
Qy 469 IleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGly 488
Db 1162 ATCGACCTGTGGCGGTGGCGGTGGAGCGGTGCTGCGGTGCGGTGCGGTGCGGTGCGG 1221
Qy 489 ValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArg 508
Db 1222 GTGCTGTGTGGGATCTCTCAAGGATCTCTGCTGCGGTGCGGTGCGGTGCGGTGCGG 1281
Qy 509 ValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerVal 528
Db 1282 CTCATCCGCGGCGCGCCAGCGCGGTGCTTGGCGAGCGGAGCGGTGCGG 1341
Qy 529 IleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeu 548
Db 1342 GGTCTACCAACATCGAGCGGTTCGAGAACCGCGAGCGGTGCGGCGGTGCTGCTCATCT 1401
Qy 549 GluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArg 568
Db 1402 GCTTCGATGCGCGCTGCTTTCGCCAACCGCGCGCATCTGCGG-----1446
Qy 569 TrpIleAspGluGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrVal 588
Db 1447 -----GAGCAGCGCGCGGCTTGTGCTCCAGCGGAGTCCCGGCTGCGGTGCTTC 1497
Qy 589 IleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGlu 608
Db 1498 GTGCTGTATGCTGCGCGGTTCGACATGGACGTCACCGCTGCGGCGGATGGGCTGGAGAG 1557
Qy 609 ValLysLysIleThrGluArgGluLeuGlnLeuValLeuValAsnProValSerGlu 628
Db 1558 CTGCGCGGTGAGTTCGAGAGAGGCGGTGCTGCTGGGCGTCCCGAGGCTCGGGCGCG 1617
Qy 629 ValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIleTyr 648
Db 1618 CTGCGCGGCTGCTGCGGCTACCGGCTGCTGAGAGCGGCTGCGGCGGAGAGTGTGCAT 1677
Qy 649 LeuThrValGluAlaVal 655
Db 1678 GCCACCGTGGAGCGGCGGT 1698

RESULT 14
US-09-902-540-768
; Sequence 768, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
```



```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4516
; LENGTH: 2832
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4516

Alignment Scores:
Pred. No.:      1,55e-68      Length:      2832
Score:          641.50       Matches:    184
Percent Similarity: 44.92%   Conservative: 143
Best Local Similarity: 25.27% Mismatches:     290
Query Match:      18.46%     Indels:      111
DB:              Gaps:       22

US-10-762-049-18 (1-680) x US-09-949-016-4516 (1-2832)

Qy      18 AsnThrIlelGluAspMetClySerValAspTyrGlnTyrProLeuGlyMetAsn--- 36
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      55 AACGTTTCACCCAGAGACTCGCTGAAGAAATGACACAGTTATCCATCTGGSATCATCTG 114

Qy      37 -----AsnPheGluArgValHisGlnValGlu 45
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db      115 GAACCTCAAGGAATCAAGTACTGACTTCAGCAATTGAGACCAATGATCAATGC--- 171

Qy      46 ValProProGlnProphePheLysSerLeu-----LysTyrSerLeuLysGluThr 63
:::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      172 -----AGACCTTATCATGATCCTTATTGAGCGGTCAAGAAATCAGATACA 219

Qy      64 PhePheProAspProLeuArgGlnPhe----LysAsn-----LysProAla--- 78
||| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      220 AACTTCAGGAGTGTGTATTAAAAAGTCGAGAAGAATTGCCAGTGCAGTCCAGGCCAAA 279

Qy      79 SerLyLeLysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLys 98
||| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      280 GCCAAATAATGATTTTAGT-----TTCCTCTGTGTTGCAAGTGCGTCCCANA 330

Qy      99 TyrThrPhe--GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeu 117
||| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      331 TAGCACCTAAGAAACAATTTTAGGGATGTGATGCAGGCTTGATGTGGGCATATTA 390

Qy      118 AlaIleProGlnGlyIleSerTyralalylsleulaalenLeuProIleLeuGlyLeu 137
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      391 TTGGTCCCCAGTCCATTGCTTATTCCTCGTGGCTGGCCAAAGAACCCTGTCTATGGTCTG 450

Qy      138 TyrSerSerPheIleProLeuIleTyralalMetMetGlySerSerArgaspLeuAla 157
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      451 TACACATCTTTTTTGGCCAGCATCATTTATTTCTCTTGGGTACCTCCCGTCACATCTCT 510

Qy      158 ValGlyThrValalaValGlySerLeuLeuMetGly----- 169
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      511 GTGGGCAATTTTGGAGTACTGTCCTTATGATTTGTTGAGACAGTGTGACGAGACTACAG 570

Qy      169 ----- 169
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db      571 AAAGCTGGCTATGACAAATGCCCATAGTCTCTCTCTTAGGAATGGTTTCAAATGGAGC 630

Qy      170 SerMetLeuSerAsnAlavalaspproanGluaspprollysLeuTyr---LeuHisLeu 188
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      631 ACATTATTAAATCATACATCAGACAGGATATGTGCACAAAAGTTGCTATGACAATATGGTT 690

Qy      189 AlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeuLeuGlyLeuPheArgLeu 208
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      691 GCGAGACCTGTAACTTTATAGCTGGAGTTTATCAGGTAGCGATGGGCTCTTTCAAGTG 750

Qy      209 GlyLeuIleValasppheLeuSerHisAlaThrIleLeuGlyPheMetGlyValala 228
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```

QY 569 -----TrrileaspGluGluGluArgIleLysAlaThrGlyGluThr--- 583
Db 1852 ATAAAGGTGGCTTGGAGAGGAGGAGAGAGAGATCAAGAGAAAAGTAGTACTCTT 1911
QY 584 -----SerLeuGlnTyr 587
Db 1912 GGTGAATCCAGGATGAAATGTCAGTGCAACTTCCCATCATCCCTTGGAGCTGCATACT 1971
QY 588 valIleIleaspMetSerAlaValGlyAsnIleaspThrSerGlyIleSerMetLeuGlu 607
Db 1972 ATAGTGAATTGACTGCAGTGCATTCATTTTATAGATACACGAGGATCCACACTGAAA 2031
QY 608 GluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProValSer 627
Db 2032 GAAGTTCGCAGAGATTATGAAGCCATTGGAATCCAGGTTCTGTGGCTCAGTGCATCC 2091
QY 628 GluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIle 647
Db 2092 ACTGTGAGGATTCCTTAACCAACGAGAGATATTGCAAAAAGGAGAGAGAAAACCTTCTC 2151
QY 648 TyrLeuThrValGluGluAlaVal 655
Db 2152 TTCTATAGTGTGTATGAAGCGATG 2175

```

Search completed: September 12, 2005, 19:19:33  
Job time : 305 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 12, 2005, 18:50:46 ; Search time 996 Seconds  
(without alignments)  
4483.658 Million cell updates/sec

Title: US-10-762-049-18  
Perfect score: 3475  
Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGWNV 680

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10762049/runat\_12092005\_120742\_7258/app\_query.fasta\_1.839  
-DB=Published Applications NA -QFMT=fastcap -SURFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -LOCAL=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10762049 @CGN 1 1 723 @runat\_12092005\_120742\_7258  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:  
22: /cgn2\_6/ptodata/1/pubpna/US10I\_NEW\_PUB.seq:  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2595.5	74.7	1977	9	US-09-938-842A-1502	Sequence 1502, Ap
2	2595.5	74.7	1977	11	US-09-938-842A-1502	Sequence 1502, Ap
3	2330.5	67.1	2372	19	US-10-437-963-38184	Sequence 38184, A
4	2036	58.6	2008	18	US-10-424-599-131116	Sequence 131116, A
5	1856	53.4	2727	20	US-10-425-115-50601	Sequence 50601, A
6	1845.5	53.1	2651	20	US-10-425-115-182207	Sequence 182207, A
7	1844.5	53.1	2366	18	US-10-424-599-65103	Sequence 65103, A
8	1811.5	52.1	2001	19	US-10-437-963-95853	Sequence 95853, A
9	1773.5	51.0	2780	19	US-10-437-963-77005	Sequence 77005, A
10	1762	50.7	2856	18	US-10-425-114-15719	Sequence 15719, A
11	1762	50.7	2713	20	US-10-425-115-75763	Sequence 75763, A
12	1734.5	49.9	2429	19	US-10-437-963-73410	Sequence 73410, A
13	1698	48.9	4390	19	US-10-437-963-77012	Sequence 77012, A
14	1676.5	48.2	1518	18	US-10-424-599-131117	Sequence 131117, A
15	1563.5	45.0	2251	19	US-10-437-963-72062	Sequence 72062, A
16	1513.5	43.6	1980	17	US-10-260-238-1066	Sequence 1066, Ap
17	1500	43.2	2336	18	US-10-424-599-52175	Sequence 52175, A
18	1460	42.0	1992	18	US-10-425-114-13333	Sequence 13333, A
19	1446.5	41.6	3252	19	US-10-437-963-78357	Sequence 78357, A
20	1436	41.3	3201	20	US-10-425-115-44188	Sequence 44188, A
21	1406	40.5	2242	20	US-10-425-115-40196	Sequence 40196, A
22	1379.5	39.7	2506	18	US-10-424-599-74393	Sequence 74393, A
23	1218	35.1	2156	18	US-10-424-599-60281	Sequence 60281, A
24	1215	35.0	2156	20	US-10-739-930-3837	Sequence 3837, Ap
25	1204	34.6	1809	18	US-10-424-599-83137	Sequence 83137, A
26	1185.5	34.1	1632	18	US-10-424-599-74389	Sequence 74389, A
27	1175.5	33.8	2213	19	US-10-437-963-76453	Sequence 76453, A
28	1134.5	32.6	2346	20	US-10-425-115-22539	Sequence 22539, A
29	1016	29.2	1303	18	US-10-424-599-139734	Sequence 139734, A
30	1010.5	29.1	1551	18	US-10-424-599-54350	Sequence 54350, A
31	1009	29.0	1247	18	US-10-424-599-130951	Sequence 130951, A
32	1002	28.8	2058	9	US-09-938-842A-717	Sequence 717, App
33	1002	28.8	2058	11	US-09-938-842A-717	Sequence 717, App
34	955.5	27.5	2388	18	US-10-437-963-92700	Sequence 92700, A
35	934	26.9	680	18	US-10-424-599-32307	Sequence 32307, A
36	923	26.6	1019	19	US-10-767-701-11661	Sequence 11661, A
37	901	25.9	1174	20	US-10-425-115-75765	Sequence 75765, A
38	900	25.9	536	18	US-10-424-599-32308	Sequence 32308, A
39	898	25.8	674	19	US-10-767-701-701	Sequence 701, App
40	892.5	25.7	1115	19	US-10-767-701-9435	Sequence 9435, Ap
41	887	25.5	1069	20	US-10-425-115-3619	Sequence 3619, Ap
42	814	23.4	1802	19	US-10-437-963-7407	Sequence 7407, Ap
43	737.5	21.2	1018	18	US-10-424-599-18417	Sequence 18417, A
44	736.5	21.2	841	18	US-10-424-599-135930	Sequence 135930, A
45	715	20.6	788	20	US-10-425-115-75767	Sequence 75767, A

ALIGNMENTS

RESULT 1  
US-09-938-842A-1502  
; Sequence 1502, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 1502  
 ; LENGTH: 1977  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-09-938-842A-1502

## Alignment Scores:

Pred. No.: 5,936-263 Length: 1977  
 Score: 2595.50 Matches: 498  
 Percent Similarity: 87.73% Conservative: 81  
 Best Local Similarity: 75.45% Mismatches: 77  
 Query Match: 74.69% Indels: 5  
 DB: 9 Gaps: 3

US-10-762-049-18 (1-680) x US-09-938-842A-1502 (1-1977)

QY 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43  
 DB 1 ATGGGACGGAGCTACACATCTCCCTCAAGGCGGAGGAGCTTCACCGCGCCACCCAC 60  
 QY 44 ---ValGluValProProGlnProPhePheLeuSerLeuLysTyrSerLeuLysGlu 62  
 DB 61 ACGGTGGAGGCTCCACAACTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120  
 QY 63 ThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82  
 DB 121 ACTCTGTTTCCAGACACCTTTTAGACAAATTAAAGAACCAAAATGTCATCAAGAAATTT 180  
 QY 83 MetLeuGlyLeuGlnPhePheProIlePheGluTyrAlaProLysTyrThrPheGln 102  
 DB 181 GGTGAGGCTCAATATCTCTCCCGATTTTCAATGGGACCGCTCAATCTCAAG 240  
 QY 103 PheLeuLysAlaAspLeuLeuAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122  
 DB 241 TTCTTCAATCAGATCTCATCGCGGAATCACCATCGCTAGTCTCGCCATCCCTCAGGCG 300  
 QY 123 IleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerPheIle 142  
 DB 301 ATCAGTTACGCCAAACTTGCTAACTTGCCCCCAATTTCTTGGGCTTTATTCGAGTTTGA 360  
 QY 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162  
 DB 361 CCGCCATTGGTATACCGGTGCTAGGAGTTCAAGGGACTTAGCGGTGGNAOAGTTGG 420  
 QY 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182  
 DB 421 GTTGGCTCTCTGTGACAGGTGGATGCTGAGCAAGAAAGTTGATGCTGAGAAAGATCCT 480  
 QY 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202  
 DB 481 AAGCTTTTACCTTCACCTGCTTTCACCGGCACCTTTTTCGCGGCGGTCTCGAAGGCTCT 540  
 QY 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGln 222  
 DB 541 CTGTGAATTTTCAAGGTTAGGTTTCATAGTGGACTTCTTATCGCATGCAACGATAGTAG 600  
 QY 222 yPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242  
 DB 601 ATTATGGGAGGAGCAGCGGTGGTGTGATCTGCAACAGCTTAAAGGTTATTTTCGGACT 660  
 QY 242 uGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnTh 262  
 DB 661 TAAACATTTTCAACAGCTCTACCGATGTTATCTCTGTGTCGCTTCGCTTCTTCTCCCAAAC 720  
 QY 262 rHisGluTyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282  
 DB 721 TCAGAGTGGAGATGGGAAGTGGCGTTCTTGAATGTGGTTTCTTATTTCTTCTCTCTC 780

QY 282 rThrArgTyrPheSerLysLysArgProArgPhePheTyrValSerAlaMetAlaProLe 302  
 DB 781 CACCAGATATTGAGCATCAAGAAACCAAAATTTCTTTGGGTGGCGGATGGCTCTCTTT 840  
 QY 302 uThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyVa 322  
 DB 841 GACCTCAGTGATTTCTTGAAGTCTCTTGGTTACTTCTCACTCAGCTGAGACATGGTGT 900  
 QY 322 lGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342  
 DB 901 TCAAGTGATAGGGACCTGAAGAAAGGTTTGAATCCACTCTCCGGTTCTGATCTCATCT 960  
 QY 342 eValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLe 362  
 DB 961 TACTTCCCTCTTACATGTCCACAGCTGTCAAAACCTGGCTCATCACTGGCATCATTTGCT 1020  
 QY 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGln 382  
 DB 1021 CGC-GNAGGAGTAGCAGTGGGAGGAGTTTTCGATGTTTCAAGAACTACACATAGACGG 1079  
 QY 382 yAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTy 402  
 DB 1080 GAACAAAGAGATGATAGCGTTTGAATGATGAACATCGTTGCTTCTTCCATCTTGTTA 1139  
 QY 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422  
 DB 1140 CCTCAACACCGGACCATTTTCAAGTCCGAGTGAACATACACGCGGTTTGAAGACCGC 1199  
 QY 422 aAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProle 442  
 DB 1200 AATGTCCAAACATAGTATGGGATTCGCTTATGTTTCACTCTCTTCTTCCACACCGCT 1259  
 QY 442 uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462  
 DB 1260 TTTTTCACCTACACACCTCGTCTCTCTGCCATCATATATCCGCAATGCTCGGACT 1319  
 QY 462 uIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCy 482  
 DB 1320 CATTGACTATCAAGTCCCATCTCTGAAAGTTGACAAAGTTTCGACTTCTCTCGTCTG 1379  
 QY 482 sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI 502  
 DB 1380 CATGAGCGCTACGTTGGGTCGTATTCGCGAGTGTAGAGATTGAGTTCGCTCGTAGCGT 1439  
 QY 502 eValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuGln 522  
 DB 1440 GCGCATATCTATAGCGAGTTTGTCTTGTGTCGAGGCGCAAACTTCGCGTGAAGGG 1499  
 QY 522 yAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa 542  
 DB 1500 AAACATACCAACACGATGATCTATAGGAACACTGAGCAGTACCCTCATCAAGAACCGT 1559  
 QY 542 lProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrIle 562  
 DB 1560 TCCTGTATTTCTCATCTTGGAGATTGATGCTCCCATCTACTTTGCTAATGCCAGTTACTT 1619  
 QY 562 uArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThrGlyGln 582  
 DB 1620 CGGTGAAGAATCATAGGTGGATTGATCAAGAGGAGAGAGAGTTAAACAATCATAGGAGA 1679  
 QY 582 uThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGln 602  
 DB 1680 GAGCAGCTTACAATATATATATCTGATGTCAGTCTGTTGATATATCGACCAAGCGG 1739  
 QY 602 yIleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValIle 622  
 DB 1740 TATTAGCATGATGGTGAATTAAGAAAGTCATTGACAGGAGCGGTTAAAGTTGCTATT 1799  
 QY 622 uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641  
 DB 1800 GTCAAAATCCAAAGGAGGTCGTGAAGAAATTAACACAGATCCAAATTTTCATCGGTGATCA 1859  
 QY 641 sLeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAs 661



Db 1860 TTTGGGCAAGAGCTGGATGTTCTTAACGGTAGGAGAGCAGTGGAGGCTTGTAGCTACAT 1919  
Qy 661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680  
Db 1920 GCTTCACAGCTTTAAACCGAACCG--GCCTCCAAAACGAGCCTTGGAAACACGTA 1974

## RESULT 2

US-09-938-842A-1502  
; Sequence 1502, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1502  
; LENGTH: 1977  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1502

## Alignment Scores:

Pred. No.: 5,93e-263 Length: 1977  
Score: 2595.50 Matches: 498  
Percent Similarity: 87.73% Conservative: 81  
Best Local Similarity: 75.45% Mismatches: 77  
Query Match: 74.69% Indels: 5  
DB: 11 Gaps: 3

US-10-762-049-18 (1-680) x US-09-938-842A-1502 (1-1977)

Qy 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43  
Db 1 ATGGGCAAGGAGGAGTACACATTCCTCAAGGAGCGGAGGAGCTTCACCGCCGCCAC 60  
Qy 44 ---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62  
Db 61 ACGGTGGAGGCTCCACAACTCAACCGTTCCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120  
Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82  
Db 121 ACTCTGTTCCAGACGACCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT 180  
Qy 83 MetLeuGlyLeuGlnPhePheProLysPheGluTrpAlaProLysTyrThrPheGln 102  
Db 181 GTGTAGGCGCTCAAAATCTTCCTCCGATTTTCGAATGGGACACGCTACATCTCAAG 240  
Qy 103 PheLeuLysAlaAspLeuLeuAlaGlyLeThrIleAlaSerLeuAlaIleProGlnGly 122  
Db 241 TTCTTCAATCAGATCTCATCGCGGAAATCACCATCGTAGTCTCGCCATCCCTCAGGCG 300  
Qy 123 IleSerTyrAlaLysLeuAlaAsnLeuProProLysLeuGlyLeuTyrSerPheIle 142  
Db 301 ATCAGTTAGCCCAAACTTGTAACCTTGCCCAATTTCTGGCCCTTTATTCGAGTTTGTGA 360  
Qy 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162  
Db 361 CCGCCCAATTGGTATACCGGTGTCAGGGAGTTCAAGGGAGCTTAGCGGTGGAAACGGTTGG 420  
Qy 163 ValGlySerLeuLeuMetGlySerMetMetLeuSerAsnAlaValAspProAsnGluAspPro 182

Db 421 GTTGGTCTCTCTGTGACAGTTCGATGCTGAGCAAAAGAGTGTGATGCTGAGAAAGATCCT 480  
Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202  
Db 481 AAGCTTTACCTTCACCTTGCTTTCACCGCCACTTTTTCGCGCGCGTTCGAAAGCTCT 540  
Qy 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleLeuG 222  
Db 541 CTTGGAAATTTTCAGGCTTAGGCTTCATAGTGACCTTCTATCGCATGCAACGATAGTAG 600  
Qy 222 yPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242  
Db 601 ATTCATGGGAGGAGCAGCGCTGGTGTGAGCTGCAACAGCTTAAGGGTATTTCGGACT 660  
Qy 242 uGluHisPheThrHisGlyAlaAspIleLeuSerValMetArgSerValPheThrGlnTh 262  
Db 661 TAAACATTTACAGACTTACCGATGTTATCTCTGTATGCTGCTGCTGCTGCTGCTGCTG 720  
Qy 262 rHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282  
Db 721 TCACGAGTGGAGATGGAAAGTGGCTTCTTCGATGTGGTTCCTTATCTTCTTCTCTCTC 780  
Qy 282 rThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLe 302  
Db 781 CACCAGATATTTGAGCATCAAGAAACCAAAATTTCTTTGGGTGGCGGCGATGGCTCTTT 840  
Qy 302 uThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVa 322  
Db 841 GACCTCAGTGATTTCTTGGAAAGTCTCTTGGTTCCTACTCACGCTGAGAGACATGGTGT 900  
Qy 322 lGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342  
Db 901 TCAAGTGATAGGGGACCTGAAGAAAGGTTGAATCCACTCTCGGTTCTGATCTCATCTT 960  
Qy 342 eValSerProTyrMetThrThrAlaValLysThrGlyLysValGlyIleLeuSerLe 362  
Db 961 TACTTCCCTTACATGTCACAGCTGTCAAAAGTGGCTCATCTGCTGCTGCTGCTGCTGCT 1020  
Qy 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGl 382  
Db 1021 CGC-GAAGGAGTAGCAGTGGGAGGAGTTTTCGATGTTCAAGAACTCAACATAGACCG 1079  
Qy 382 yAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCys 402  
Db 1080 GAACAAAGAGATGATAGCGTTTGGAAATGATGAACATCGTTGCTTCTTCCATCTTGTGA 1139  
Qy 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422  
Db 1140 CCTCAACCGGACCATTTTCAAGGTGGGAGTGAACACTACAACGGGGTTCGAAGACCGC 1199  
Qy 422 alaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLe 442  
Db 1200 AATGTCCAACTAGTATGAGGCTTGGGTTATGTTTACACTCTTCTTCTCACACCGCT 1259  
Qy 442 uPheHisTyrThrProLeuValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462  
Db 1260 TTTTCACTACACACCATCTGCTGCTCTCTGCTCATCATCATCGCAATGCTCGACT 1319  
Qy 462 uIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCy 482  
Db 1320 CATTCGATATCAAGCTGCGCATCTCTGGAAGTGAACAAGTTCGACTTCTCTCTCTG 1379  
Qy 482 sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAla 502  
Db 1380 CATGAGCCCTACGTTGGGCTCGTATTCGCGAGTGTAGAGATTGAGATCGTCTCGCTGT 1439  
Qy 502 eValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuG 522  
Db 1440 GCGGATATCTATAGCGAGGTGTTGCTGTTGCTGCGAGGCCCAAACTGCGGTGAGGG 1499  
Qy 522 yAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa 542

```
Db      1500 AAACATACCAACAGCATGATCTATAGGAACACTGAGCAGTACCCTCATCAAGAACCGT 1559
Qy      542 lProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrIle 562
Db      1560 TCCTGGTATCTCATCTTGAGGATTGATGCTCCCATCTACTTTGCTAAATGCCAGTTACTT 1619
Qy      562 uArgGluArgIleThrArgTTrpIleAspGluGluGluArgIleIleAlaThrGlyG1 582
Db      1620 GCGTGAAGAAATCATAAAGTGATGATGAAGAGGAAGAGAGATTAAACAATACGAGAGA 1679
Qy      582 uThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerG1 602
Db      1680 GAGCAGCTTACATATATATATATATGATGTCAGTGTGTTGTAATATCGACACAGCGG 1739
Qy      602 yIleSerMetLeuGluGluValIleIleAspMetSerAlaValGluArgGluLeuGlnLeuVal 622
Db      1740 TATTAGCATGATGTTGGAATAATGAAGAACTCATTCACAGGAGCGTTAAAGTTGGTATT 1799
Qy      622 uValAsnProValSerGluValMetIleAspMetSerAlaValGluArgGluLeuGlnLeuVal 641
Db      1800 GTCAATCCAAAGGAGAGAGGTCGTGAAGAAATTTAACAGATCCAAATTCATCGGTGATCA 1859
Qy      641 sLeuGlyLysLysTTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAs 661
Db      1860 TTTGGCAAAGAGTGGATGTTCTTAACGGTAGAGNAGACGTTGGAGGCTTTGTAGTACAT 1919
Qy      661 nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTTrpAsnAsnVal 680
Db      1920 GCTTCACAGTTTAAACCGAACCG--GCCTCCAAAGAGAGCGCTTGGAAACAACGTA 1974

RESULT 3
US-10-437-963-38184
; Sequence 38184, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 38184
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRP4530_41845C.1
US-10-437-963-38184

Alignment Scores:
Pred. No.:      7,67e-235      Length:      2372
Score:          2330.50      Matches:    441
Percent Similarity: 83.84%      Conservative: 83
Best Local Similarity: 70.56%      Mismatches: 100
Query Match:      67.06%      Indels:    1
DB:              19          Gaps:      1

US-10-762-049-18 (1-680) x US-10-437-963-38184 (1-2372)

Qy      43 GlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db      149 AGGGTCCGATGTCGGCGCGAGAGCGCTTCTGGAGACGCTCGGGGGGAACATGAGAGGAG 208
Qy      63 ThrPhePheProAspAspProLeuArgGlnPhe--LysAsnLysProAlaSerLysLys 81
Db      1289 GCGATGTCGAACGTGATCATGTGCGGTGCGGTGATGATCACGCTGCTGTTCTTTCGCGCG 1348
```

```
Db      209 ACATTCTCTCGGAGACGCCGTTTCAGGGTGGTGGCGGGAGCGCGGGTGGCGGGCGGC 268
Qy      82 PheMetLeuGlyLeuGlnPhePheProIlePheGluTTrpAlaProLysTyrThrPhe 101
Db      269 CGCGCGCGCGCGCTCCCGTACGTGTTCCTCGTTCATGAGAGTGGCGCGCGTTCGTACACCCCTC 328
Qy      102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121
Db      329 GGCACCTTCAAGTCCGACCTCATCGCGGATACCATTCAGCAGCTCGCCATCCCCAG 388
Qy      122 GlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPhe 141
Db      389 GGCATCAGTCAGCAAGCTCGCAACCTCCTCCCGTCTCGGCTCTATTTCAGACTTC 448
Qy      142 IleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal 161
Db      449 GTGCGCGCGCTGGTGTACGCGATATGGGAGCTCGAGGGACCTGGCGGTGGGAGCGGTG 508
Qy      162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db      509 GCGGTGGCGTCGCTGCTCATCGGTTCGATCTGACGAGGAGGTGTCGCGCGCGGAGGAC 568
Qy      182 ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db      569 CCGGCGCTGTACCTGCACGCTCACCGCCACCTTCTTCGCGGCGTGTTCACGCGC 628
Qy      202 AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221
Db      629 CTGCTGGGGGTCTTGAGGCTGGGGTTCATCGTGGACTTCCTGTGCACGCCACCATCGTC 688
Qy      222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db      689 GGGTTCATGGCGCGCGCGCCACCCGCTGTGTGTGCTGACGAGCTCAAGGGCATGTTCGCGC 748
Qy      242 LeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln 261
Db      749 CTCGACCACTTCACCACCGCCACCGACCTCGTCTCCGTTCATGAGTCCGTCTTCTCCCAA 808
Qy      262 ThrHisGluTTrpArgTTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeu 281
Db      809 ACCACCTCTCGCGATGGGAGCGTCTCATGGCTCGCGCTCTCTCTTCTTCTCTCTC 868
Qy      282 SerThrArgTTrpPheSerLysLysArgProArgPhePheTTrpValSerAlaMetAlaPro 301
Db      869 ATCACCCCTCTCTCAGCAAGAGGAGGCAAGGTCTTCTGGGTATCTGCAGCTCGGCCA 928
Qy      302 LeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGly 321
Db      929 TTGGCGTCTGTTCATCATCGGAGCGCTCTGCTGTACCTCACACATGCTGAAAACCATGCGC 988
Qy      322 ValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerSerLeuThrAsnLeuVal 341
Db      989 ATTCAAGTGATTGGTTTACCTGAAGAAGGGCGCTGAACCCACCGCTCTGCCACAGCTGAAC 1048
Qy      342 PheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSer 361
Db      1049 TTTCGTCCCGTATCATGATGCTGGCCCTGAAGACCGGAGATCATCATTGGCGGTCAATTGCC 1108
Qy      362 LeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAsp 381
Db      1109 CTCGCTGAGGGGATTGCCGTAGGAAGGAGCTTTGCAATGTTCAAGAAATTACCACATCGAC 1168
Qy      382 GlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCys 401
Db      1169 GGAAACAAGGAGATGATCGCGTTTGGAAACGATGAACATTGTGGGATCACTCACCTCTTTCG 1228
Qy      402 TyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThr 421
Db      1229 TACCTCACACCGGTCGCTTCTCGAGGTGCGCGCTCACTCAACAGCCCGGCTGCAAGAGC 1288
Qy      422 AlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrPro 441
Db      1289 GCGATGTCGAACGTGATCATGTGCGGTGCGGTGATGATCACGCTGCTGTTCTTTCGCGCG 1348
```

```
QY 442 LeuPheHisTyrThrProLeuValValLeuSerAlaIleValSerAlaMetLeuGly 461
Db 1349 CTGTTCCATACACGCGCTGGTGGTGTCTGGCGATCATCATGTGGCGATGTGGGG 1408
QY 462 LeuIleAspTyrGluAlaIleHisLeuPheLysValAspLysPheAspPheValVal 481
Db 1409 CTGATCGATACCGCGCCCGCTCCACCTGTGGCAGTGGCAAGGTGACTTCTCGCTC 1468
QY 482 CysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAla 501
Db 1469 TGCCTCGCGCCCTACCTCGCGTGTCTTCCGAGCGTGGAGATCGGCTCGTGTGCGCC 1528
QY 502 IleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeu 521
Db 1529 GTCGGATCTCCATCTCCGCGTGTCTGTTCTGGCGCGCCGAGGAGCAGCGTGTCTC 1588
QY 522 GlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHis 541
Db 1589 GGGAAACATCCCAACACGATGATCTACCGCGCGATGGACAGTACACCCCGCGCAGAGG 1648
QY 542 ValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyr 561
Db 1649 GTCCCCGGGTGTCTGTCTCGCGTGTCTGCTGTCGCGCGATGGACAGTACACCCCGCGCAGAGG 1708
QY 562 LeuArgGluArgIleThrArgTyrPheGlyGluGluGluArgIleLysAlaThrGly 581
Db 1709 CTGCGTGAGAGGATCCGCGGTGGATCGACGAGGAGGACCATGTCGCAAGGAGGCG 1768
QY 582 GluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSer 601
Db 1769 GAGATGGCATTCAGTACGTCTGTCTCGACATGGGTGCGGTGCGAAGCATCGACACCGAGC 1828
QY 602 GlyIleSerMetLeuGluValLysIleThrGluArgArgGluLeuGlnLeuVal 621
Db 1829 GGGACGAGCATGCTGGATGAATCAGGAAGACCTGGACAGAGGCGGCTTCAGATTGTG 1888
QY 622 LeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHis 641
Db 1889 TTGGCAACCCGGGAGCGAGATCATGAGAGAGTTGGACAGCTCCDAGGTGCTTGAGGCG 1948
QY 642 LeuGlyLysTyrPheTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661
Db 1949 ATCGGCGATGAGTGGATCTTCCCAACGCTGGCGAGGCGGTGGCGAGTGGCGACTTCGTG 2008
QY 662 LeuArgAlaSerLys 666
Db 2009 ATGCACTCGCAGAG 2023
```

## RESULT 4

```
US-10-424-599-131116
; Sequence 131116, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131116
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89404C.1
US-10-424-599-131116
Alignment Scores:
```

```
Pred. No.: 7.31e-204 Length: 2008
Score: 2036.00 Matches: 403
Percent Similarity: 97.3% Conservative: 8
Best Local Similarity: 95.50% Mismatches: 9
Query Match: 58.59% Indels: 2
DB: 18 Gaps: 0
US-10-762-049-18 (1-680) x US-10-424-599-131116 (1-2008)
QY 261 GlnThrHisGluTyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
Db 3 CAAAGCATGAGTGGAGGTGTGAAGCGCTGTGTGGGTGTGTGTGTGTATTTTCTTCCTC 62
QY 281 LeuSerThrArgTyrPheSerLysArgProArgPhePheThrPheValSerAlaMetAla 300
Db 63 CTCGTCAACAGATACTTTCAGCAACAGCAGACCAAGTTCTTCTGGGTGTGCACAATGGCA 122
QY 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db 123 CCACCTGACGTCGCTTATATTTGGGAAGTCTGCTGTGTCTATGTGCACACACGCTGAGAACAT 182
QY 321 GlyValGlu--ValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnL 340
Db 183 GGAGTTCAAGTGTATAGAGAAATTTGAAGAGGGTTTGAATCCACCATCATCTGACAAATC 242
QY 340 euValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleI 360
Db 243 TGGTATTTGTGACGCTTACATGACCAAGCTGTCAAAACAGGCAATTTGCTGTGGCATTA 302
QY 360 LeuSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnI 380
Db 303 TATCTCTTGGCGAAGAAATAGCAGTGGGAAGAGCTTTTGCAATGTATATAAAATTCATAA 362
QY 380 LeuAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrS 400
Db 363 TTGATGGCAACAAAGAGATGATAGCTATTGGAGCCTGAACTAGTGTGTTTCTTCACT 422
QY 400 erCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysL 420
Db 423 CTTCGTCTCTCACACTGACCACTTTTCGCGTTTCAGCTGTGAACATATAACGCTGGATGCA 482
QY 420 ysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuT 440
Db 483 AGACAGCAGCATCCAAACATCATATATGTCAATGTGCAATGTGCAATGTGTTATTTCTTAA 542
QY 440 hrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetL 460
Db 543 CACCAATTTGTCATTACACTCTCCCTGGTGGTGTATCAGCCATTTATTTGATCTGCAATGC 602
QY 460 euGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheV 480
Db 603 TTGGCTCTCATAGATTATGAGCAGCGCATCTCATTTTAAGGTCCGACAAATTTGACTTTG 662
QY 480 alValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValI 500
Db 663 TCGTGTGCATGAGTGCATACGTGTGGCGTGTCTTTGGCAGTGTGCAAAATTTGGCTTAGTCA 722
QY 500 leAlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheV 520
Db 723 TAGCTATTGGCGATATCTGTACTTTCGGGTACTTCTATTTATTGCAAGGCCAAGGACATTG 782
QY 520 alLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaL 540
Db 783 TTTTGGGAAACATTCCAAATTCGTGTGATATACAGAAATTTGTGAGCAGCTATCTCAATGCAA 842
QY 540 ysHisValProGlyMetLeuLeuGluIleAspAlaProIleTyrPheAlaAsnAlaS 560
Db 843 AACATGTTCTCGAATGCTAATTTCTAGAGATTGATGCAACCAATTTTACTTTTGTGCAATGCAA 902
QY 560 erTyrLeuArgGluArgIleThrArgTyrPheLeuAspGluGluGluArgIleLysAlaT 580
Db 903 GCTATTTAAGAGAAAGGATCAAGATGGATTGATGGATGAGGATGAGAAAGAAATCAAGCTA 962
```



```
Db 1623 GCATGCATGGAGCATTTTGGAGTCATATTTTCATCGGTGAGTATGCTTGTCTCAIT 1682
Qy 501 AlaileValIleSerValLeuArgValLeuPheIleAlaLeuProArgThrPheVal 520
Db 1683 GCGGTTGCAATCTCTTGTAAATTTCTTCCAAAGTAAACAGCGGCAAGACAGTTTAA 1742
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db 1743 CTTGGAAACCTTCCACGAACACTATATACAGGAATGTAGAACAGTATCTGTATGCTACC 1802
Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db 1803 AAGGTTCCAGCGGTGCTGATTGTAGAGTGGACTCAGCTATATATCTTCAAAACTCTAAC 1862
Qy 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
Db 1863 TATGTTAAAGAGAGATCTCGAGGTGGCTAAGAGATGAGGAGGAGCAACACAGACCAG 1922
Qy 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db 1923 AAGTTAAACAAAACCTGATTGTTCTTAATGTGTGACTGTCTCTCTTAATTGATTCGACACA 1982
Qy 601 SerGlyIleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeu 620
Db 1983 AGTGGAAATCCATGCTTTGGAGGAGTTGGCGAAAGCTCTTGAAAAACGCAAAATTCAGCTG 2042
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
Db 2043 GTTCTTACCAATCCCGGGCGCGGGTGGATCCAGAAAGCTCCGGCTCAGCGAAATTCACGGAC 2102
Qy 641 HisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPhe 660
Db 2103 ATGATTGGTGAGACAAACATATTTCTCAGCGTGGCGACGCTGTG----- 2147
Qy 661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsn 678
Db 2148 -----NAGAAATTTGCTCTTAAGATGTCGATAATGCTGATGGAGT 2189

RESULT 6
US-10-425-115-182207
; Sequence 182207, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 182207
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_97750C.1
US-10-425-115-182207

Alignment Scores:
Pred. No.: 1.5e-183 Length: 2651
Score: 1845.50 Matches: 355
Percent Similarity: 74.41% Conservative: 116
Best Local Similarity: 56.08% Mismatches: 157
Query Match: 53.11% Indels: 5
DB: 20 Gaps: 3

US-10-762-049-18 (1-680) x US-10-425-115-182207 (1-2651)
Qy 41 ValHisGlnValGluValProGlnProGlnProPhePheLysSerLeuLysTyrSerLeu 60
```

```
Db 518 GTGCACAGGTGGCGCGCGCGCGCGAGCACCGCGAGCAAGATGAAGGTGAGGGTG 577
Qy 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
Db 578 AAGGAGACCTTCTTCCCCAGCACCGTTCCCGGGGCTTCAAGGCGGCGCGCGGGGAGC 637
Qy 81 LysPheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThr 100
Db 638 CAGTGGCTCATGCGCGCTCAGGTACCTTTCCTCCCATCTCTGGACTGGGTGCGAGCTACTCC 697
Qy 101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
Db 698 TTGTGCTCTTCAAGTCCGACCTCGTCGCGCGCTCACCATTCGCCAGCTCGCAATTCCT 757
Qy 121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSer 140
Db 758 CAGGGCATTAGCTACGCAAGCTGGCAAGCTTGCCTTCCCTATTCCTGGGTGATTTCGAGC 817
Qy 141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
Db 818 TTGTGCGCGCGCATGGGTACCGGTGCTGGGGGAGCTCCCGTGACCTGCGGTGGCGCGCG 877
Qy 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
Db 878 GTGTGATCTCGTCTGATCATGCGGTCCATGCGCGAGCGCTGCGCGAGCCCGCACCGCG 937
Qy 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
Db 938 GAGCGAGCGCTGTCTCGAGCTGGCGCTTCCCTCCACCTGTTCGCGGGGTGGTGCAG 997
Qy 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
Db 998 GCTCTCGTGGGATCTCTAGGCTCGGTTCGTCATCGACTTCTGTCCAAGCGAGCGCTG 1057
Qy 221 IleGlyPheMetGlyIleAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeu 240
Db 1058 GTGGGTTCATGCGCGCGCGCCCATCATCGTGGCGCTGCAGCAGCTCAAGGCGCTGCTG 1117
Qy 241 GlyLeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThr 260
Db 1118 GGATCGTCTCCATTCACCAACCGAGATGGGCATCGTCCCATGATGGCCCTCCGCTTCCAC 1177
Qy 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeu 280
Db 1178 CACACGAGCGAGTGTGCTGGGAGGATCCTCATGGGGCTGTGCTTCTCTGCTTCTCTG 1237
Qy 281 LeuSerThrArgTyrPheSerLysArgProArgPhePheTrpValSerAlaMetAla 300
Db 1238 CTGTGCGGAGGATGTGAGCATCAGATGGCCAAAGCTTTTCTGGGTTCGCGGTGCGCG 1297
Qy 301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db 1298 CCCCTGGCATCGTCAACATCTCGACGCTGTGTTTCTCTTCAAAAGCTCAGAACCAT 1357
Qy 321 GlyValGluValIleGlyLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
Db 1358 GGATCAGCATCATTTGGGCGAGCTCAAGTGGCGCTGAATCGCCCTCTGGGGCAAGCTC 1417
Qy 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIle 360
Db 1418 CTGTTTGACACGGCGTATTAGGCGCTCACCATGAGAGCTGGCTTCTCACCGGAATCATC 1477
Qy 361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380
Db 1478 TCACTGACGGAAGGATAGCGGTGGTAGAACATTTTGCCTCACTCAAGGACTACCAAGATA 1537
Qy 381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
Db 1538 GATGGAAACAAGGAGATGATGCGCATGAGGTGATGATGTTGTGGGTCTCTGCACATCA 1597
Qy 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
```

```
Db 1598 TGCTAGCTAAACAGGTCGGTTCTCCCGCTCTGCTGTATAAACACCAACACCGCGCTGCAAG 1657
Qy 421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrIleLeuPheLeuThr 440
Db 1658 ACTGCCATCTCCAACGTCATGATGGCGCTGACTGTGTGTCACGCTGCTGTCTCTCATG 1717
Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db 1718 CCACGTGTTGCTGTACACACCAACGTTGCTCGGAGCGATCATATCGCCGCGGTGATC 1777
Qy 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspPheVal 480
Db 1778 GGCCTGATCGATTCTCCCGCGGTGTACCACTCTGGAGATGACAGATGATTTCTG 1837
Qy 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValIleIleGlyLeuValIle 500
Db 1838 GTGTGGTGTGGCGGTTTCCCGCGGTGCATCTTCATCTCAGTCCAAGAGCGCTTGGATA 1897
Qy 501 AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal 520
Db 1898 GCGGTTGGTATATCTATATTTAGGGTGTGATCGAGATCACCAGGCCCAAGATGATGGTT 1957
Qy 521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db 1958 CAAAGGAACATCAAGGGGACTGATATTTACAGAGACCTGCATCACTACAGAGGCCCA 2017
Qy 541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db 2018 AGAGTTTCTGGGTTCTTGATCTTGGCCATTGAAGCACCGCATAAACTTGGCCCAACTCCAAC 2077
Qy 561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThr 580
Db 2078 TACCTGAATCAAGAGGATTAAGATGGATA-----GAGGAAGAATCTTTTGAACAGGAT 2131
Qy 581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db 2132 AACATACCTGAACCTCAATTCATPAATCTTGGATCTGTGACGTGTTCTCTCAATTGACACA 2191
Qy 601 SerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeu 620
Db 2192 AGTGGCATAGCGTCTCTCATTAAGAAATCAATAGAGAAACGTCGTCTGGAGCTT 2251
Qy 621 ValLeuValAsnProValSerGluValMetLysLeuAsnLys---SerLysPheGln 639
Db 2252 GTGCTTGTCAATCCAACTGGAGAAGTCATGGAGAAAATACAACGTCGCAACGAGGCTGAA 2311
Qy 640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 659
Db 2312 AACTATTTTAGCCAGATTGCTGTGTATCTGACCACCTGGCGAAGCAATC-----GCTTCA 2365
Qy 660 PheAsnLeuArgAlaSerLysThrAsnProLysLysAsp 672
Db 2366 CTTTCTGCACTTGCCCAAGATGACAAAACCCCTTAATGAT 2404
```

## RESULT 7

```
US-10-424-599-65103
; Sequence 65103, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 65103
; LENGTH: 2366
; TYPE: DNA
; ORGANISM: Glycine max
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29800C.1
US-10-424-599-65103
```

```
Alignment Scores:
Pred. No.: 1598-183 Length: 2366
Score: 1844.50 Matches: 367
Percent Similarity: 73.64% Conservative: 133
Best Local Similarity: 54.05% Mismatches: 165
Query Match: 53.08% Indels: 15
DB: 18 Gaps: 5
```

US-10-762-049-18 (1-680) x US-10-424-599-65103 (1-2366)

```
Qy 8 SerTyrIleThrHisIleCysLeuLeuArgAsnThrIleIleGluAspMetGlySerVal 27
Db 73 ACTTACATCATCACACCACTTTCTCTTATA-AGTTCCTGTCTCTCTCAAAATGGAGACTAAT 131
Qy 28 Asp-----TyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
Db 132 AATGATGATGACTATGATTTCTCATTCGATAGAGATGTCC-----ATGGAAGTGCACCA 185
Qy 44 ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGluThr 63
Db 186 GTTGTGCCACCACTTCACAGAGACCCCTTCAGAAACTCAAGGTAGGCTCAAGGAAACT 245
Qy 64 PhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPheMet 83
Db 246 TTCTTCCCTGATGATCCTCTGCGCAATTCAAGGACCAACCTCTTAAGAGAAAAACTGATC 305
Qy 84 LeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPheGlnPhe 103
Db 306 CTTGAGGCTCAATATGTGTCCTTATCTTCAATGGGTCTTAAGTATAATCTTCAAACTC 365
Qy 104 LeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIle 123
Db 366 TTCAATCTGACCTGTGTTCTGGCCCTCACTATTGTAGCTTGGACATCCCGCAGGGAATG 425
Qy 124 SerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIlePro 143
Db 426 AGTTATGCTTAGCTTGCAGAGTCTTCTCCAATTTTAGGACTTTATCTTAGTTTGTGTTCA 485
Qy 144 ProLeuIleTyrAlaMetMetGlySerArgAspLeuAlaValGlyThrValAlaVal 163
Db 486 CCACGTGCTATGCTGTCTTGGAGAGCTCAAGAGACCTTGCAGCTTGGAGCTGTTCTATT 545
Qy 164 GlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLys 183
Db 546 GCTTCTCTTGATGGATGCCATGTTGCATCAGGAAGTGTCTCCCAACACAGATCCAATT 605
Qy 184 LeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaLeu 203
Db 606 CTGTTCTTTCAGCTAGCTTTCACATTAATTTTGCCTGCTCTCTTTCAAGCTTTTGCTT 665
Qy 204 GlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPhe 223
Db 666 GGAATCCTTAAGGCTAGGCTTCAATAATGATTTTCTATCTTAAGCCCACTTATTTGGTTC 725
Qy 224 MetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGlu 243
Db 726 ATGGCTGGAGCTGCTATTATTGCTCTCAGTGCACAGCTCAAGAGCTGCTTGGAAATCACA 785
Qy 244 HisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHis 263
Db 786 CATTTCACTTAATCAGATGGTCTGATTCCTGTATTGACTTCTGTTTTCACAAATATACAT 845
Qy 264 GluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThr 283
Db 846 GAGTGGTCAATGGCAACCAATATTGATGGGATTTGCTTCTTGTTACTACTACTATTAGCA 905
Qy 284 ArgTyrPheSerLysLysArgProArgPheThrValSerAlaMetAlaProLeuThr 303
Db 906 AGACACGTTAGCATTAAGGAACCAAACTATTCTGGGTCTCAGCTGGAGCTCCTCTTATG 965
```



```
Qy 304 SerValIleLeuGlySerLeuLeuValTyPheThrHisAlaGluTyHisGlyValGlu 323
Db 966 TGTGTCAATCTCTACCCCTCTGGTTTTCGAATTAAGGCTCAAAATCATGGCATCAGT 1025
Qy 324 ValIleGluLeuLeuGlySerLeuLeuAspProSerLeuThrAsnLeuValPheVal 343
Db 1026 CGGATTGGAAATTTGCAACAGGAATAAATCTCCATCATCGAATATGTTGCTCTTTCAT 1085
Qy 344 SerProTyRMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAla 363
Db 1086 GGAAGTCACCTGAGCCTAGTTATGAAACACAGGGCTTATCACCGGATTTTGTCCCTAAC 1145
Qy 364 -GluGlyIleAlaValGlyArgSerPheAlaMetTyRLeuAsnTyRAsnIleAspGlyAs 383
Db 1146 GGAAGGTATTGTCAGTAGGAAGACATTTGCAGCTCTCAAAACATACAAAGTGATGGAAA 1205
Qy 383 nLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyRLe 403
Db 1206 TAAAGGAATGATGGCAATGGGTTTATGAATGTTTGGTCTCTTCACCTTCCTGCTATGT 1265
Qy 403 uThrThrGlyProPheSerArgSerAlaValAsnTyRAsnAlaGlyCysIleThrAlaAl 423
Db 1266 TACAACAGGTGCTTCTCTCGGTGAGCAGTTAACACATGTCAGCGCGCAAAAACAGCTGT 1325
Qy 423 aSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPh 443
Db 1326 GTCAAATGATGATGTCGTGACAGTCATGGTGACACTCTTTTCTCATGCCATGTT 1385
Qy 443 eHisTyRThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuI 463
Db 1386 TCAATACACGCCATAATGCTGCTGGCGCAATCATAGTCACAGCAGTAATGGCCTCAT 1445
Qy 463 eAspTyRGlulAlaAlaIleHisLeuPheTySValAspLysPheAspPheValValCysMe 483
Db 1446 CGATCTCCCGCTGCTGTGAACATTTGGAAGATCGACAAATTCGATTTTGTGTGATAT 1505
Qy 483 tSerAlaTyRILEGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleVa 503
Db 1506 GATGCTTCTTAGTGTTCTTTTTCATCTCTGTCGAAGGCGCTGCTCTGCTGTGG 1565
Qy 503 lIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAs 523
Db 1566 ATTATCAACTTTGAAGATACTTTGCAAAATTACGAGACCGCAAAACAGTGATGTTGGGAA 1625
Qy 523 nIleProAsnSerValIleTyRArgAsnValGluHisTyRGlulAsnAlaIleHisValPr 543
Db 1626 GATACCGAGGAACAGACATATATAGAAATCTTGATCAATACAGGAAGCTGTGAGAAATACC 1685
Qy 543 oGlyMetLeuIleLeuGluIleAspAlaProIleTyRPhleAlaAsnAlaSerTyRLeuAr 563
Db 1686 TGAATTTCTATTATTAAGCATTGAGGCTCCCATCAATTTTGCCAAATCATCATATCTCAA 1745
Qy 563 gGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGluTh 583
Db 1746 CGAGAGAACGTTAAGATGGATTGAAGAAGAAGAACAAACATAAAGAA-----CAATT 1799
Qy 583 rSerLeuGlnTyRValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyI 603
Db 1800 AACCCCTTCGATTTCTAGTATTGAAATGTGAGCTGTGAGTGCTGTTGACACAAAGTGAAT 1859
Qy 603 eSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValLeuVa 623
Db 1860 CTCACCTTTCAAGAGATTGAAACCAACACTGGAAGAAAGAGGTTGTGAGCTGTGTGTGT 1919
Qy 623 lAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn----HisLe 642
Db 1920 CAATCTCTTGTGAGTGCATAGAAAAGCTTAAAAAAGCAGATGAAGCTAATGATTTTCAT 1979
Qy 642 uGlyLysLysTyRTrpIleTyRLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLe 662
Db 1980 AAGAGCAGATAACCTTTTCTTGACGTGGAGAGGCTGTAGCT-----TCACT 2027
```

```
Qy 662 uArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db 2028 TTCTTCAGCAATGAAGGCCAATCATCAACCATTACAGAAGGGGCACACACAATT 2082
RESULT 8
US-10-437-963-95853
; Sequence 95853, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 95853
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94005C.1
US-10-437-963-95853
```

```
Alignment Scores:
Pred. No.: 3 7e-180 Length: 2001
Score: 1811.50 Matches: 340
Percent Similarity: 73.10% Conservative: 122
Best Local Similarity: 53.80% Mismatches: 167
Query Match: 52.13% Indels: 3
DB: 19 Gaps: 3
US-10-762-049-18 (1-680) x US-10-437-963-95853 (1-2001)
```

```
Qy 43 GinValGluValProProGlnProPhePheLysSerLeuLysTyRSerLeuGlyGlu 62
Db 106 GAGGTGAACCTGTCCGGGGCGCGCGCTTCGCGGAGAACGTGTGGTCGACCTCGCGGAG 165
Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProLaserLysLysPhe 82
Db 166 ACGTTCCTTCCCGACGACCCGTTCCGCGGGTTTCGCGCGCTGCCCGCCGCGCGCGCG 225
Qy 83 MetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyRThrPheGln 102
Db 226 TGGTGGCGGTCAGTACTTCGTCGCGGCTGGACTGGTCCGCGCTACGCGCTCGAC 285
Qy 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 286 AAGTTCAGTTCCGCTCTCTCCGCGCATCACCATCGCCAGCTCGCCATCCGCGAGGCG 345
Qy 123 lIleSerTyRAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyRSerSerPheIle 142
Db 346 ATCAGCTACGCGCGCTCGCCAACTCCCTCCAAATCATCGGCTCTTATTCGAGCTTCGTG 405
Qy 143 ProProLeuIleTyRAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 406 CCGCGGTGATGATGACGGGTGTCGGAGCTCAACACACTGGCGGTGGGACGGTGGCG 465
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
Db 466 GCGGCGTCTGCTCTGCGGCTCATCATCGAGACGAGGAGTGTGCGCGCGGACGAGACCCG 525
Qy 183 LysLeuTyRLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 526 CAGCTGTACTCAGCTCTTCTTACACCGCGCGCTTCTTCCAGCGCGCTCTTCCAGACCGCG 585
```



QY 203 LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGly 222  
DB 586 CTAGGGCTCTTATAGGTAGGGCTAATAGTGGATTCTCTCGCGTCGACGATCACCGGG 645  
QY 223 PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeu 242  
DB 646 TTCATGGGGCGGACGGCGATGATCATCATCTACTGCAGCAGTTCAAGGGCCCTCTGGGGATG 705  
QY 243 GluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThr 262  
DB 706 AAGCACTTCACCAACAGACAGACATCATCTCGTCTCGCATCTCCACCTACCATTAACCGG 765  
QY 263 HisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuSer 282  
DB 766 CACGAGTGGAAATGGCAGAGTCAGTTCTCGGCATATGCTCTCTTGTCTGTGATGTCG 825  
QY 283 ThrArgTyrPheSerIlyValArgProArgPhePheThrValSerAlaMetAlaProLeu 302  
DB 826 AGTAAGCACTGGAAGAAGAACTGCCAAAGTTGTTCTGGGTGTCGGCCATTCGGCCATTC 885  
QY 303 ThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluIlyHisGlyVal 322  
DB 886 ATGTTGTGTCATCGAGGCACTTCTCGCTTCTTCTTCAAGGGTCAAGCAGCATGGAAAT 945  
QY 323 GluValIleGlyLeuLysIlyGlyLeuAsnProProSerLeuThrAsnLeuValPhe 342  
DB 946 CCGATAGTCGGTAACCTGAATAAGGGATCAATCCCTCTGTCATTCGCGCAGCTAACGTTTC 1005  
QY 343 ValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeu 362  
DB 1006 CAGTCACGACATGAAGCGGAGTGAAGCTGGTCTCTTGTCTGGGATCTTAGCAGTCG 1065  
QY 363 AlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGly 382  
DB 1066 GCAGAAGGAATACCCGTCGCGCGAAGCTTGGCGATGCTCAAGAACGAGCAGATCGACGCG 1125  
QY 383 AsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr 402  
DB 1126 AACCAAGGAGATGATCGCTTGTGTATCATGAACATCGCCGGTCTTCTCACCTCCTGCTAC 1185  
QY 403 LeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAla 422  
DB 1186 CTCACACAGCGCGCTTCTCAAGTCGGGGTGAACCTTCACGCCGGTGCAGACGCGCG 1245  
QY 423 AlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeu 442  
DB 1246 ATGTGCAACGTGGTATGCGGTGTCATCATCTGCTGCTGCTGCTTCTTGGCACCCTTC 1305  
QY 443 PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu 462  
DB 1306 TTCAAGTACACCCCGTGGTGGCGCTCTCTCCATCATCTGTCGTCGCCATGATCGGGCTC 1365  
QY 463 IleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCys 482  
DB 1366 GTCAGGTCAAGAGATTCGCCACCTCTACAGGTGCACAGTTTCGACTTCTGCACTGCG 1425  
QY 483 MetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle 502  
DB 1426 ATGGTCGCTTCTCGGGGTGCTTCTTCCACCATGTCGAGCGCTTCAGCGCATCCGTA 1485  
QY 503 ValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGly 522  
DB 1486 GGCTTGTGATGTCAGGGGTGCTGTGTAGTCGGCCAGCGCTGCCACCTGCAAGCTCGGG 1545  
QY 523 AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVal 542  
DB 1546 AACATGACGGAGCGAGACCTTCGCGGAGCTGAGACATACCCCAAGCGAAGACATC 1605  
QY 543 ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu 562  
DB 1606 CCTGGCATCTCTGTCAGCTCGGCTCTCCCATCTACTTCTCAATCGGGTTACCTG 1665  
QY 563 ArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGlu 582

DB 1666 CGAGAAAGATTTTCAGATGGTGAAGATGAGGATAACCTGTGCAAGAGCGTCGGGCAC 1725  
QY 583 ThrSerLeuGlnTyrValIleAspMetSerAlaValGlyAsnIleAspThrSerGly 602  
DB 1726 --GATCTGCAATACTTGGTCTTCTGATCTGGTGGTGTCACTTCTGTGCAAACTCTGGA 1782  
QY 603 IleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValLeu 622  
DB 1783 GTTGGATGCTACTAGAAGTACACAAGAGCCTCGAACGAAGAGGATCACGATAGCTCTG 1842  
QY 623 ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641  
DB 1843 ACGAATCCGAGGTAGAGTACAGAGAGCTGGTCTCTGGATAGCTCAGGGACATC 1902  
QY 642 LeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661  
DB 1903 TTAGGGGATGAGTGGTCTTCTGACGCTCAAGGACGCCATCAGCGCGTGTGATACGCG 1962  
QY 662 LeuArgAlaSerLysThrAsnProLysLysAspGlu 673  
DB 1963 CTCGAGATATCCAGA---AATAAGGAGGAAGACGAA 1995  
RESULT 9  
US-10-437-963-77005  
; Sequence 77005, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221) B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 77005  
; LENGTH: 2780  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_76946C.1  
US-10-437-963-77005  
Alignment Scores:  
Pred. No.: 6,51e-176 Length: 2780  
Score: 1773.50 Matches: 328  
Percent Similarity: 73.1% Conservative: 128  
Best Local Similarity: 52.6% Mismatches: 166  
Query Match: 51.04% Indels: 1  
DB: 19 Gaps: 1  
US-10-762-049-18 (1-680) x US-10-437-963-77005 (1-2780)  
QY 34 GlyMetAsnAsnPheGluArgValHisGlnValGluValProProGlnProPhePhe 53  
DB 433 GGCCATCATCATCATCATCATCATCGGCCACAGAGTTCGAGTTCCTCCACCAAGAGAGTTCATC 492  
QY 54 LysSerLeuLysTyrSerLeuLysGluThrPhePheProAspAspProLeuArgGlnPhe 73  
DB 493 GACGAGTTCACCGACGCGCGTGAAGGAGAGCTTCTTTCGCCGACGACCCCGTTCGGGAGTAC 552  
QY 74 LysAsnLysProAlaSerLysLysPheMetLeuGlyLeuGlnPhePheProIlePhe 93  
DB 553 AAGGACCCAGCCGATGTCAAAGAGAGTGTGATCAGCCCTGCCAAACTTCTTCCCGGTGCTG 612  
QY 94 GluTrpAlaProLysTyrThrPheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThr 113





```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_169123C.1
US-10-425-115-75763

Alignment Scores:
Pred. No.: 1,02e-174 Length: 2713
Score: 1762.00 Matches: 334
Percent Similarity: 73.23% Conservative: 120
Best Local Similarity: 53.87% Mismatches: 162
Query Match: 50.71% Indels: 4
DB: 20 Gaps: 2

US-10-762-049-18 (1-680) x US-10-425-115-75763 (1-2713)

Qy 41 ValHieGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
Db 356 CTGCACAAAGGTGCTGCTGCGGAGCGCGGAGCGCGGAGCGGCTGCGGAGCGCCCTG 415

Qy 61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
Db 416 GCCGAGGTGTTCTTCCGCGAGCGCGCTGCACCAAGTTCAGAAACCAAGTCGTCGGCGCG 475

Qy 81 LysPheMetLeuGlyLeuGlnPhePheProLlePhePheLysTrpAlaProLysTyrThr 100
Db 476 CGCCTCGTGTGGCGGTGCACACTACTTCTCCCATCTTCCAGTGGGGTCCGCTACAGC 535

Qy 101 PheGlnPheLeuLysAlaAspLeuLleAlaGlyLeThrLleAlaSerLeuAllePro 120
Db 536 CCGCGCTCTCGCTGCGCTCGACCTGCTGCGCGGCTCACCATTCCAGCTCGGCATCCCG 595

Qy 121 GlnGlyLeuSerTyrAlaLysLeuAlaAsnLeuProProLleLeuGlyLeuTyrSer 140
Db 596 CAGGGATCAGCTACGCCAGCTCGCCACCTGCCCCCAATCGTTGGCTATATTCAGC 655

Qy 141 PheLeuProLeuLleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
Db 656 TTGCTGCGCGCGCTCATACGGCGTGTGGGAGCTCGCGGACCTGCGCGTGGGGCGG 715

Qy 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAnGlu 180
Db 716 GTGTCTCATCGCTCGCTGTGATGGGGTCCATCTCCGGAGCGCGTGTGCGCGGAGCAG 775

Qy 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
Db 776 CAGCGCTCTCTACCTGAGCTGGGCTTCACCGCCACCTTCTTCGCGCGGTCTTCCAG 835

Qy 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuLleValAspPheLeuSerHisAlaThrLle 220
Db 836 GCGTCCCTGGGATTCCTCAGGCTGGGCTTCATCGTGGACTTCTGTCCAAGCGCGCTG 895

Qy 221 IleGlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerLleLeu 240
Db 896 ACGGGCTTCATGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 955

Qy 241 GlyLeuGluHisPheThrHisGlyAlaAspLleSerValMetArgSerValPheThr 260
Db 956 GGCATCTCCCATCTACCTCCACATGGGATTCCTCGAGTCTATGCGCTCGCTCGCTCAAC 1015

Qy 261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheLlePhePheLeu 280
Db 1016 CGCCACGACGAGTGGAGTGGACAGCATGTCATGGGCTCGGCTTCTCGCCATCCTC 1075

Qy 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
Db 1076 CTCCTCAGCGCGCAATCAGCGCGCAGGAAACCCAAAGCTTTTCTGGGTATCAGCAGGTGCT 1135

Qy 301 ProLeuThrSerValLleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db 1136 CCCTGGGTGGGTGATCATCTCCACCATCTCTCTCTCATCTGCG-----AAATCCCCC 1189

Qy 321 GlyValclnValLleGlyGluLeuLysGlyLeuAsnProProSerLeuThrAsnLeu 340
Db 1189
```

## RESULT 12

```
US-10-437-963-73410/c
; Sequence 73410, Application US/10437963
; Publication NO. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
```

```
Db 1190 AGCATCAGTGTATTATGGCATCTCCCGAGGGAGTGAACCCCTCTTCCGCGAACATGCTC 1249
Qy 341 ValPheValSerProTyrMetThrThrAlaValLysThrGlyLleValValGlyLleLle 360
Db 1250 AGCTTCAGCGGCTCTTAATGTGGCGTACGATCAAAACCGGATCATGACAGCATCTCGT 1309
Qy 361 SerLeuAlaGluGlyLleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnLle 380
Db 1310 TCTTTAACAGAGAGGATCGCAGTGGCAGGACCTTCGCGTCCATCAAACTACCAAGGTG 1369
Qy 381 AspGlyAsnLysGluMetLleAlaLleGlyThrMetAsnValValGlySerPheThrSer 400
Db 1370 GACGGAAACAGGATGATGCGCTCGGCTGATGAACATGCGCGGCTCTCCGCGCTCC 1429
Qy 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
Db 1430 TGCTAGTGACGACGCGGCTCTTCTCCGCTCGGCGGTGAATACAGCGCGGCTCGCAGG 1489
Qy 421 ThrAlaAlaSerAsnLleLleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db 1490 ACGGCGCTGTCCAACTGATGCGCGCGGCTGTGCTGACGCTGCTGTCTCTCATG 1549
Qy 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaLleLleValSerAlaMetLeu 460
Db 1550 CGCTGTTCCTACATACACCCGACGATCTCTGCGCGGATCATCATCACGCGGTGGTG 1609
Qy 461 GlyLeuLleAspTyrGluAlaAlaLleHisLeuPheLysValAspLysPheAspPheVal 480
Db 1610 GGGCTGTGGAGCTGCGCGCGCGCGGCTGTGAGAGTGGAGACAGCTGACACTTCTCTG 1669
Qy 481 ValCysMetSerAlaTyrLleGlyValValPheGlySerValGluLleGlyLeuValLle 500
Db 1670 GGTGCGTGGCGGCTCTCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1729
Qy 501 AlaLleValLleSerValLeuArgValLeuPheLleAlaArgProArgThrPheVal 520
Db 1730 GCGCTGCGGCTCTCTCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1789
Qy 521 LeuGlyAsnLleProAsnSerValLleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db 1790 GAGGCGCTGCTCCCGGGGACGACAGCTACCGCGCGCTGCGCGCTGCGCGCGCGCTC 1849
Qy 541 HisValProGlyMetLeuLleLeuGluLleAspAlaProLleTyrPheAlaAsnAlaSer 560
Db 1850 CGCGTGGCGGCTCTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1909
Qy 561 TyrLeuArgGluArgLleThrArgTrpLleAspGluGluGluGluArgLysAlaThr 580
Db 1910 TACCTGTGGAGCGGTATGCTGCTACCTCCGCGACGAGGAGCGCGCTCAAGTCC 1969
Qy 581 GlyGluThrSerLeuGlnTyrValLleLleAspMetSerAlaValGlyAsnLleAspThr 600
Db 1970 AACCACTCCATCCGATGCTGCTCTCGACATGGGCGCGCTCGCGCGATTCGACAGC 2029
Qy 601 SerGlyLleSerMetLeuGluValLysLleThrGluArgGluGluGluGlnLeu 620
Db 2030 ACGGCTTAGACGCGCTGCTCGAGCTCAAGAAAGTCTCGACAAAGAAACATCGAGCTG 2089
Qy 621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
Db 2090 GTGCTTGGCAACCGGTGGGTGCGTGGCGGAGAGATGTTCAACTCGCGGTGGCGGAG 2149
Qy 641 HisLeuGlyLysLysTrpLleTyrLeuThrValGluAlaVal-----GlyAlaCys 658
Db 2150 AGCTTGGGTGCGGCGCGCTCTTCTTACGCTAGCGGAGGCGCTCGCGCGGCGGTGC 2209
```

```

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 73410
; LENGTH: 2429
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73694C.1
US-10-437-963-73410

```

```

Alignment Scores:
Pred. No.: 6,85e-172 Length: 2429
Score: 1734.50 Matches: 333
Percent Similarity: 70.93% Conservative: 116
Best Local Similarity: 52.61% Mismatches: 163
Query Match: 49.91% Indels: 21
DB: 19 Gaps: 3

```

US-10-762-049-18 (1-680) x US-10-437-963-73410 (1-2429)

```

QY 41 ValHisGlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeu 60
DB 2261 CTCGCAAGGTCTCGGTGCGGAGGAGGTTCGAGCGGCAAGGCGCTGAGCGCGGTG 2202

QY 61 LysGluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
DB 2201 GCGGAGGTCTTCTCCCGAGCAGCCGCTCCACAGTTCAAGAACACGTCGTCGCGCGG 2142

QY 81 LysPheMetLeuGlyLeuGlnPhePhePheProLlePheGluTrpAlaProLysTyrThr 100
DB 2141 CGCGTGGTCTCGCGCTGAGTACTTCTTCCCACTCTCCATTTGGGGCTCCGACTACAGC 2082

QY 101 PheGlnPheLysAlaAspLeuLleAlaGlyLleThrLleAlaSerLeuAlaLlePro 120
DB 2081 CTCGCGCTCTCGCTCCGAC----- 2061

QY 121 GlnGlySerTyrAlaLysLeuAlaAsnLeuProProLleLeuGlyLeuTyrSerSer 140
DB 2060 ---GGAAATCAGTACGCCAAGCTCGCAACCTGCCTCCCAATCATTTGGACTATATTCGAGC 2004

QY 141 PheLleProProLeuLleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
DB 2003 TTCTGTCGCGCTGATCTACTCTGTTGGTGGTAGCTCGCGGATCTGGCTGTAGGCGCG 1944

QY 161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
DB 1943 GTGTGATAGCGTCTGATGATGGGTTCGATGCTCCGAGCGGCGGTGTCGCGGACCCAG 1884

QY 181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
DB 1883 GAGCCCATCTCTACCTCCAGCTCGCTTCACCTCCACTTCTTCCGCGCGGTCTTCCAA 1824

QY 201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuLleValAspPheLysSerHisAlaThrLle 220
DB 1823 GCCTCCCTCGGCTTCTCCGCGTGGGTTCATCGTGGATTTCTTGTGCAAGGCGAGCGCTG 1764

QY 221 IleGlyPheMetGlyAlaAlaThrValValCysLeuGlnGlnLysSerLleLeu 240
DB 1763 ACGGGATTATGCGGCGCGCGGATCATCGTGTGCTGCAGCAGCTCAAGGGATTCGTC 1704

QY 241 GlyLeuGluHisPheThrHisGlyAlaAspLleIleSerValMetArgSerValPheThr 260
DB 1703 GGGATCATCCATTTCAGTCGCGAGATGGGGTTCGTCAGGTGATGCTCCGCTCTTCAG 1644

```

```

QY 261 GlnThrHisGluTrpArgTyrGluSerAlaValLeuGlyCysValPheLlePhePheLeu 280
DB 1643 CACCACGACGAGTGGCGTGGCAGACCATCTCACTGGCGTCGCGCTTCTCGCGCTCTC 1584

QY 281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
DB 1583 CTCACACACGCCACATCAGCGCCAGGACCCAAAGCTTTTCTGGGTGTGACGAGTCT 1524

QY 301 ProLeuThrSerValLleLeuGlySerSerLeuValTyrPheThrHisAlaGluLysHis 320
DB 1523 CCACGTGACATCAGTGTATCTCTACCATCATCTCGTTTCGTGAGCAAGCT-----CAT 1470

QY 321 GlyValGluValLleGlyCylLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
DB 1469 GGCATCAGTGTATTGGCGATCTCCCAAGGATTTGAACCTCTCTTACAGCAATATGCTG 1410

QY 341 ValPheValSerProTyrMetThrAlaValLysThrGlyLleValValGlyLleLle 360
DB 1409 ACCTTCAGTGGCTCTACGTAGGACTGGCTTAAACACAGGATTTATGACTGGCATCTA 1350

QY 361 SerLeuAlaGluGlyLleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnLle 380
DB 1349 TCTCTCACTGAAGGATAGCAGTAGGACGACATTTGCATCCATCAACAACTACCAAGT 1290

QY 381 AspGlyAsnLysGluMetLleAlaLleGlyThrMetAsnValValGlySerPheThrSer 400
DB 1289 GATGGAAACAAGGAGATGATGCCATTGTTGTCATGAACATGGCTGCTCATGTGCCCTCC 1230

QY 401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
DB 1229 TGCTATGTACACACAGGATCTGTTCTGAGGTCTGGCGGTGAACACAGCGCGGGTGCAG 1170

QY 421 ThrAlaAlaSerAsnLleLleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
DB 1169 ACGCGGTGTGCAACATCGTATGCGGTGCGGTGCTGTTGAGCGCTGCTGTCTCTGATG 1110

QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaLleLleValSerAlaMetLeu 460
DB 1109 CGGTGTTTCCACTACACCCCAAGGTATCTCTCGCGGATCATCATCACCCTGCTGATC 1050

QY 461 GlyLeuLleAspTyrGluAlaAlaLleHisLeuPheLysValAspLysPheAspPheVal 480
DB 1049 GGCCTCATCGAGTCCGGGCGCGCCAGGCTGTGGAAGTTCGACAACTCGACTTCCTC 990

QY 481 ValCysMetSerAlaTyrLleGlyValValPheGlySerValGluLleGlyLeuValLle 500
DB 989 GCCTGATGCGCGCTTCTCGCGTCTCTCGTCTCGTCCGATCGGCTCGCCATC 930

QY 501 AlaLleValLleSerValLeuArgValLeuPheLleAlaAspProArgThrPheVal 520
DB 929 GCGGTGCGCATCTCCCTTCAAGATCTCTCTCAGGTCCACCCGCCCAACATGTGTGTC 870

QY 521 LeuGlyAsnLleProAsnSerValLleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
DB 869 AAGGCGTCTGCGCGGACCGGAGCTACCGAGCATGGCGCATCAGGAGGCCATG 810

QY 541 HisValProGlyMetLeuLleGluLleAspAlaProLleTyrPheAlaAsnAlaSer 560
DB 809 CGTGTGCGCTCTCTCTCGTGGCTCGAGTCCGCTCGCATCTACTTCCCAACTCCATG 750

QY 561 TyrLeuArgGluArgLleThrArgTrpLleAspGluGluGluArgLleLysAlaThr 580
DB 749 TACCTCGCGAGAGGATCATGAGTTTCTCGGAGGAGAGACGAGCGGCGGCCCAAGTGC 690

QY 581 GlyGluThrSerLeuGlnTyrValLleLleAspMetSerAlaValGlyAsnLleAspThr 600
DB 689 AACAGTCCCTGTGATGATCATCTCTCGACATGAGTGTGTGTCAGCGATTCGACAGC 630

QY 601 SerGlyLleSerMetLeuGluValLysLysLleThrGluArgArgGluGlnLeu 620
DB 629 AGTGGCTCGATTCGACGAGCTGAAGAGGTGCTGGAAAAAAGAAACATCGAGCTT 570

```

QY 621 ValLeuValAsnProValSerGluValMetLysLeuAsnLysSerLysPheGlnAsn 640  
DB 569 GTGCTGGCCAAACCGGTTGGATCGGTGACCGAGAGGCTGTACAACCTCGGTGGTCGCGAAG 510  
QY 641 HisLeuGlyLysLysTrpLleThrValGluGluAlaValGlyAlaCysAsnPhe 660  
DB 509 ACCTTCGGCTCAGACCGCGGTGTTCTTACAGCGTCGCGAGGCGGTGCGGCGCGC----- 456  
QY 661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGlu 673  
DB 455 -----CCGCACAAACACAGCCCTGCACAGAGCAG 426

## RESULT 13

US-10-437-963-77012  
; Sequence 77012, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 77012  
; LENGTH: 4390  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_76952C.1  
US-10-437-963-77012

## Alignment Scores:

Pred. No.: 1,29e-167 Length: 4390  
Score: 1698.00 Matches: 328  
Percent Similarity: 72.61% Conservative: 120  
Best Local Similarity: 53.16% Mismatches: 164  
Query Match: 48.86% Indels: 6  
DB: 19 Gaps: 2

US-10-762-049-18 (1-680) x US-10-437-963-77012 (1-4390)

QY 42 HisGlnValGluValProProGlnProPheLysSerLeuLysTyrSerLeuLys 61  
DB 503 TACAATGTGGCGCGCGCGAGAGAGAACCTCTCGCGGAGTTCCGCGCAGCGTGAAG 562  
QY 62 GluThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLys 81  
DB 563 GAGACGTTCTTCCGACGAGCGATGCGCGGTACAGGACCAGCGAGGTCCAGGAAG 622  
QY 82 PheMetLeuGlyLeuGlnPhePheProLlePheGluTrpAlaProLysTyrThrPhe 101  
DB 623 CTATGGCTCGCCTTGCAGCACGCTTCCCGGTGTTCCGAATGGGCGACACAATACACCCCTC 682  
QY 102 GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121  
DB 683 GCCAAGTTACGGGCGACCTCATTTGCGGCTCACCTTGCACGCTCTGCATACCTCAG 742  
QY 122 GlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerPhe 141  
DB 743 GACATCGGCTACGGAAGCTTGTAACCTGCCACAGAGATTGGGCTGCACAGTAGCTTC 802  
QY 142 IleProProLeuLleTyrAlaMetMetGlySerArgAspLeuAlaValGlyThrVal 161  
DB 803 GTCCCGCCATTGATATACGCTCTGTGGGCGACCTCAAGGGAGCTAGCAATGGGTCCAGTG 862

QY 162 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181  
DB 863 GCGGTCTCATCTACTGCTGCTGCTCTCTCTCCAGGAGGAGTACTCAAAAGAAGAAC 922  
QY 182 ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201  
DB 923 CGCTAGATTACAGACGGCTCGCTTCACAGGAGCTTCTTTGCGAGGAGTCACACAGGCG 982  
QY 202 AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221  
DB 983 GCGCTGGGTTCGTCAGGCTAGGTTTCATCATAGCGTTCCTCTCTCATCTGCTGCCATCATC 1042  
QY 222 GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241  
DB 1043 GGATTTCATGGCGCGCGCGCCCATCACCATTGCTCTTCAGCAGCTTAAAGGCTTCTCTTGA 1102  
QY 242 LeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln 261  
DB 1103 ATTGCAAACTTACCACAGAGACTGACATCATCTCGGTGATGAAATCAGTCTGGGGAAT 1162  
QY 262 ThrHis-----GluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePh 279  
DB 1163 GTTCACCATGGGCAATGGAACTGGCAGACATATTGATCGGAGCATCATTTTGGCATT 1222  
QY 279 eLeuLeuSerThrArgTyrPheSerLysLysArgProArgPhePheThrValSerAlaMe 299  
DB 1223 CCTCCTGGT-TGCCAA-----GGCAAGAAGAACAAAGAGCTTCTCTGGGTCCAGCAAT 1275  
QY 299 tAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLys 319  
DB 1276 TGCACCACTCATTTCCGGTGATCATTTCACTTTTGTCTCTATCATCACTCTCGTGCACAA 1335  
QY 319 shiSGlyValGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAs 339  
DB 1336 ACAGCGTGGCAATCGTAAACAGTCAAGAAAGCATCAATCACCTTCAGTAGCT 1395  
QY 339 nLeuValPheValSerProTyrMetThrAlaValLysThrGlyIleValValGlyL 359  
DB 1396 GATATTTTCTACTGGCCCATCTGCTCAAAGGATTCAAATTTGAGTAGTAGTGAAT 1455  
QY 359 eIleSerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAs 379  
DB 1456 GATAAGCCTTACGGAAGCGATTGCGAGTTGGAAGAACATTTGCTGGATTGAACGATTACCA 1515  
QY 379 nIleAspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheTh 399  
DB 1516 GATAGATGGGAACAAAGAAATGTTGGCTCTAGAACCATGAATGTTGGTTCGTTCAATGAC 1575  
QY 399 rSerCysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCy 419  
DB 1576 GTCTTGTATATAGCCACAGGTGGTTTTCACGATCAGCAGTCAATTTGCTGGTGGAGG 1635  
QY 419 sLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLe 439  
DB 1636 TAAACACCAATATGTCATATTTGTTATGTTATGTTACCTAGTAGTATTGTTGCACTCTCTG 1695  
QY 439 uThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMe 459  
DB 1696 CACTCCATTTGTTCAAGTACAGCCCAATGCCACCATTTCTTCCATCATCATATCAGCAT 1755  
QY 459 tLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPh 479  
DB 1756 GCTTGGCCTATTTGATCTTTGAATCAGCTACCTTATCTGGAAGTTGATAAGTTGACATT 1815  
QY 479 eValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVa 499  
DB 1816 CATGCGGTCTTGGGCGCATTCCTTGGAGTAATAATTTTCATCTGTGGAGTATGGCTTGTCT 1875  
QY 499 tIleAlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPh 519  
DB 1876 CATTCGGGTGTAATATCACTAATCAAAGTTCTGCTCCATGTAACACGCGCCAGGACAGC 1935  
QY 519 eValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAl 539









```
QY      580 ThrGlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAsp 599
Db      1696 GTCCGCGGCGAGGACCTGCTACGTCGTCTCGACATCGCGGCGGTGACGGCGATCGAC 1755
QY      600 ThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGln 619
Db      1756 AGCCCCGGGATCGAGATGCTGCGGGAGGTGCACGGCGAGCTGGAGAGGAAGGGGATGAAG 1815
QY      620 LeuValLeuValAsnProValSerGluValMetLysLysLysLeuAsnLysSerLysPheGln 639
Db      1816 ATGGCGGTACGACACCGAGGATGGCGGTGGCGGAGAGCTGTGCTGTCCGGGCTCGCG 1875
QY      640 AsnHisLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsn 659
Db      1876 GAGCTGTCGGCGAGAGCTGGATGTTCTCTCCAAATGGCGACGCTTGGCGCGGTGCCGG 1935
QY      660 PheAsnLeuArgAlaSerLys 666
Db      1936 TACAGCTCCAGGGCTCCAAG 1956
```

Search completed: September 12, 2005, 21:40:45  
Job time : 1049 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 31, 2005, 09:44:04 ; Search time 8026 Seconds  
(without alignments)  
4105.351 Million cell updates/sec

Title: US-10-762-049-18

Perfect score: 3475

Sequence: 1 HELARTLSYTHICLLRNTI.....NLRASKTNPKDTEGNNV 680

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Rgapop 6.0 , Rgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US10762049/runat\_30082005\_121245\_5972/app\_query.fasta\_1.839  
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10762049 @CGN 1 1 4545 @runat\_30082005\_121245\_5972 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEOQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	3475	100.0	2449	6 AR477464
2	2642.5	76.0	2045	8 D89631
3	2641.5	76.0	2266	8 AY099873
4	2597.5	74.7	2193	8 BNA581745

5	2595.5	74.7	1977	6	AX506807	Sequence
6	2595.5	74.7	1977	6	AX651748	Sequence
7	2330.5	67.1	1974	6	AX653455	Sequence
8	2330.5	67.1	2327	8	AK066932	Oryza sat
9	2314	66.6	2037	8	AB004060	Arabidops
10	2274.5	65.5	2219	8	AJ601439	Brassica
11	2164	62.3	1981	6	AR477457	Sequence
12	2026.5	58.3	123209	8	ATF4F15	Arabidops
13	2020.5	58.1	60000	8	AF049236	Arabidops
14	2012.5	57.9	7012	8	AB012048	Arabidops
15	1901	54.7	2067	6	AR477463	Sequence
16	1898.5	54.6	1896	6	AX651700	Sequence
17	1894.5	54.5	1917	8	AB023423	Arabidops
18	1892.5	54.5	2470	8	AF347614	Arabidops
19	1875	54.0	2256	8	AB054645	Arabidops
20	1853	53.3	2640	8	AF355602	Zea mays
21	1847	53.2	144217	8	AC139748	Medicago
22	1844.5	53.1	2279	6	AR477456	Sequence
23	1841.5	53.0	2159	8	AB042322	Arabidops
24	1836.5	52.8	2181	8	AK104831	Oryza sat
25	1833.5	52.7	2238	8	AK072809	Oryza sat
26	1830.5	52.7	2299	8	AJ704374	Brassica
27	1817	52.3	1940	6	AX654768	Sequence
28	1811	52.1	2167	8	BNA311388	Brassica
29	1809.5	52.1	2166	8	AJ633706	Brassica
30	1807.5	52.0	1971	6	BD175021	Gene enco
31	1807.5	52.0	1971	8	AB049624	Arabidops
32	1807.5	52.0	2000	6	AX653563	Sequence
33	1799	51.8	2310	8	BT009249	Triticum
34	1799	51.8	2311	6	AR477465	Sequence
35	1796.5	51.7	2248	8	AF332959	Lycopersi
36	1793.5	51.6	2808	8	AK121195	Oryza sat
37	1791.5	51.6	2327	8	AF347613	Lycopersi
38	1789	51.5	2310	8	AB061739	Arabidops
39	1777.5	51.2	1989	8	AX653919	Sequence
40	1777.5	51.2	2791	8	AK103007	Oryza sat
41	1773.5	51.0	1989	8	AF493792	Oryza sat
42	1771.5	51.0	2494	8	AK067270	Oryza sat
43	1770.5	50.9	2262	8	BNA416460	Brassica
44	1767.5	50.9	2439	8	AF309643	Solanum t
45	1760.5	50.7	2442	8	HVHVT1	Horduem vul

ALIGNMENTS

RESULT 1  
AR477464  
LOCUS AR477464 2449 bp DNA linear PAT 14-MAY-2004  
DEFINITION Sequence 17 from patent US 6696292.  
ACCESSION AR477464  
VERSION AR477464.1 GI:47234922  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2449)  
AUTHORS Allen,S.M., Falco,S.C. and Thorpe,C.J.  
TITLE Genes encoding sulfate assimilation proteins  
JOURNAL Patent: US 6696292-A 17 24-FEB-2004;  
FEATURES  
source Location/Qualifiers  
1..2449  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 5.03e-268 Length: 2449  
Score: 3475.00 Matches: 680  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-762-049-18 (1-680) x AR477464 (1-2449)	
Qy	1 HisGluLeuAlaArgThrLeuSerTyrIleThrHisIleCysLeuLeuArgAsnThrIle 20
Db	2 CACGAGTAGCTCGCACATTAAGTTATATTAACATATTTGCTTGTAGAAATACATAT 61
Qy	21 IleGluAspMetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArg 40
Db	62 ATTGAAGATATGGGAGTAGATTATGAGTACCTTTGGCGATGAACAACATTTTGAGA 121
Qy	41 ValHisGlnValGluValProProGlnProPhePhePheLysSerLeu 60
Db	122 GTGCACCAAGTCGAGGTTCCACCGCCACAGCCGTTTTCAGGTCTCTAAAGTAGTACTTTG 181
Qy	61 LysGluThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLys 80
Db	182 AAGGAGACTTCTCTCCCTGATGACCTTTGAGGAGTTCAAGAACAGCCAGCTTCAAG 241
Qy	81 LysPheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThr 100
Db	242 AAGTTTCATGCTTGGCCTTCAGTCTCTCTCCCATTTTCGAATGGGCTCCCAATACACC 301
Qy	101 PheGlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIlePro 120
Db	302 TTTTCAGTCTTTGAAGCTGACCTACAGCTGGCATCACCATCGCTAGCTTGGCATTCTCT 361
Qy	121 GlnGlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSer 140
Db	362 CAGGGCATCAGTTATGCCAAGCTGCCAAGCTCCCACTCCCTCCCAATCTTGACATATTCGAGC 421
Qy	141 PheIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThr 160
Db	422 TTTTATACCACTATGATTTATCGATGATGGGTAGTCGAGGAGTTTGGCAGTGGGGACT 481
Qy	161 ValAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGlu 180
Db	482 GTGGCGGTTGGATCGCTTCTGATGGGTTCGATGTGAGTAATGCCGTGTATCCCAATGAA 541
Qy	181 AspProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGln 200
Db	542 GACCCAAAGCTTTACCTCCACCTGGCTTTCACAGCTACATTTATGCTGGTGTTCAG 601
Qy	201 AlaAlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIle 220
Db	602 GCTGCTCTGGGCTGTGTTAGTTGGGGTTCGATCGTGGAATTTCTGTGCATGCAACCAT 661
Qy	221 IleGlyPheMetGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeu 240
Db	662 ATAGGGTTTCATGGGAGAGCAGCCCGTGTGTGTGTCGAGCAACTAAATCGATCTT 721
Qy	241 GlyLeuGluHisPheThrHisGlyAlaAspIleSerValMetArgSerValPheThr 260
Db	722 GGCTTTGAGCATTTACCCATGAGCTGATATCATATCAGTGATGGCTCTGTTTTACCC 781
Qy	261 GlnThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeu 280
Db	782 CAAACTCATGAGTGGAGTGGGAAAGTGTGTGTAGGATGTCTTCATTTCTTCTCTC 841
Qy	281 LeuSerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAla 300
Db	842 CTTAGCACAAAGATACCTCAGCAAAAACGACCAAGGTTTTTTTGGGTGTCGCAATGGCG 901
Qy	301 ProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHis 320
Db	902 CCATTGACGTCCTTATATGGGAAGTCTCTTGGTTTATTTCACTCACCCCGGAGGAC 961
Qy	321 GlyValGluValIleGlyLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeu 340
Db	962 GGTGTTGAAGTGTAGAGAACTGAAGAAGGGTTGAATCCACCACTCACTCAAAATCTG 1021
Qy	341 ValPheValSerProTyrMetThrAlaValLysThrGlyIleValValGlyIleIle 360
Db	1022 GTATTTGTGTCGCTTACATGACTACAGCTGTCAAAACTGGCATTTGCTGTGGCATCAT 1081
RESULT 2	
D89631	LOCUS
D89631	DEFINITION
D89631	ACCESSION
D89631.1	VERSION
ATST1; sulfate transporter.	KEYWORDS
Arabidopsis thaliana (thale cress)	SOURCE
D89631 2045 bp mRNA linear PLN 14-FEB-2004	
Arabidopsis thaliana mRNA for sulfate transporter, complete cds.	
D89631 GI:2285884	
ATST1; sulfate transporter.	
Arabidopsis thaliana (thale cress)	
361 SerLeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle 380	
Db	1082 TCACCTGGCGAAGAAATAGCAGTAGGAAGAGCTTTGCAATGTATAAAAAATTACAATATT 1141
Qy	381 AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer 400
Db	1142 GATGGCAACAAAGAGATGATAGCTATTGGGACCATGAACTGATAGTTGGTTCTTTACCTCT 1201
Qy	401 CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys 420
Db	1202 TGTCTACCTCACAAACAGGACCATTTTGGCGTTCGGCTGTGAACCTATAACGCTGGATGCAAG 1261
Qy	421 ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db	1262 ACAGCAGCTTCCAAACATATATATGTCACCTGCAGTAATGTTGACATTTGTTATTCCTGACA 1321
Qy	441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db	1322 CCCTTGTTCATACACTCCCTCGTGTGTCTATCAGCTTATTATCGTATCTGCAATGCTT 1381
Qy	461 GlyLeuIleAspTyrGluAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db	1382 GGACTCATAGATTATGAAGCAGCCATCCATCTATTATAAGGTTGACAAATTTGACCTTTGTG 1441
Qy	481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db	1442 GTGTGCATGAGTGATACATTTGGCGTGTCTTTGGCAGCTGTGGAATTTGGCTTAGTCATA 1501
Qy	501 AlaIleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheVal 520
Db	1502 GCTATTGTAATATCTGACTTCGGGTACTTCTATTATTGCAAGGCCAAGGACATTCGTT 1561
Qy	521 LeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys 540
Db	1562 TTGGGCACATTTCCAAATTTCTGTGATATACCGAAATTTGTGAGCCTATCCAAATGCCAAA 1621
Qy	541 HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer 560
Db	1622 CATGTTCTGGAAATGCTAAATTTAGAGATTGATGCACCAATTTACTTTGCCAATGCCAGC 1681
Qy	561 TyrLeuArgGluArgIleThrArgTrpIleAspGluGluGluGluArgIleLysAlaThr 580
Db	1682 TATTTTAAGAGAAGGATCACAAAGTTCGATTGATGAAGAAGAAAGAAATTAAGCTACA 1741
Qy	581 GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr 600
Db	1742 GGGGAGACTAGTTTGCAGTATGTTATAATGTATATGAGTGTCTGTGGAAACATTGATACA 1801
Qy	601 SerGlyIleSerMetLeuGluValLysIleThrGluArgArgGluLeuGlnLeu 620
Db	1802 AGTGGAAATAAGTATGCTTGAAGAGGTGAAGAAGATTACAGAGAGAAGAGAGCTACAGCTT 1861
Qy	621 ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn 640
Db	1862 GTTTTGGTCAATCTCTGAAGTGAAGTGAAGAACTGAACAAATCGAAGTTCAAAAAT 1921
Qy	641 HisLeuGlyLysLysTyrIleTyrLeuThrValGluAlaValGlyAlaCysAsnPhe 660
Db	1922 CATTTAGGGAAGAAATGAGATCTATCTGACTGTGTGAAGAGGCGGTTGGAGCATGCAACTTC 1981
Qy	661 AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTyrAsnAsnVal 680
Db	1982 ANCTACGTGCAAGCAAAACGAAACCCAAAGAAAGATGAACAGAGGGTTGGAACTATGTG 2041

ORGANISM	Arabidopsis thaliana	Db	273	TTCTTCAATCAGATCTCATCGCGGAAATCACCATCGTAGTCTCCCATCTCCCTCAGGGC	332
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.	Qy	123	IleSerTyrAlaIysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIle	142
AUTHORS	Yamaguchi, Y., Nakamura, T., Harada, E., Koizumi, N. and Sano, H.	Db	333	ATCAGTTACGCCAAACTTCTTAACCTTGGCCCCCAATCTTGGCCCTTTATTCGAGTTTGTGA	392
TITLE	The electronic Plant Gene Register	Qy	143	ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla	162
MEDLINE	Plant Physiol. 113 (4), 1463-1465 (1997)	Db	393	CGCCATTGGTATACCGGTGTAGGAGTTCAGGGACTTCAAGGACTTACGCGTGGAAACGGTTGCG	452
REFERENCE	2 (bases 1 to 2045)	Qy	163	ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro	182
AUTHORS	Yamaguchi, Y.	Db	453	GTTCGCTCTCTGTGTGACAGTGGATGCTGAGCAAGAAAGTTGATCTGAGAAAGATCCT	512
TITLE	Direct Submission	Qy	183	LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla	202
JOURNAL	Submitted (29-NOV-1996) Yube Yamaguchi, Nara Institute of Science and Technology; 8916-5 Takayama, Ikoma, Nara 630-01, Japan (E-mail:yu-yama@bmailgate.aist-nara.ac.jp, Tel:07437-2-5652, Fax:07437-2-5659)	Db	513	AAGCTTTTCACTTCACTTCCACCTGCTTTCACCGCCACTTTTTCGCGGCTTCTCGAAGCCTCT	572
FEATURES	Location/Qualifiers	Qy	203	LeuGlyLeuPheArgLeuGlyLeuIleValAppPheLeuSerHisAlaThrIleLeGly	222
source	1..2045	Db	573	CTTGGAAATTTTCAGGTTCAGGTTCATAGTGGACTTTCTATCGCATGCAACGATAGTAGGA	632
gene	/organism="Arabidopsis thaliana"	Qy	223	PheMetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeuGlyLeu	242
CDS	/mol_type="mRNA"	Db	633	TTCATGGGAGGACGACGCGGTGGTGTGAGTCTGCAACAGACTTAAGGATATTTTCGGACTT	692
	/db_xref="taxon:3702"	Qy	243	GluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThr	262
	/issue_type="above-ground parts"	Db	693	AAACATTTACAGACTCTACCGATGTTATCTCTGTCTGTCATGCGTTCCTCCCAAACT	752
	/clone_lib="lambda ZAP II"	Qy	263	HisGluTyrArgTyrGlnSerAlaValLeuGlyCysValPheIlePhePheLeuSer	282
	/ecotype="Landsberg erecta"	Db	753	CACGAGTGAGATGGGAAAGTGGCGTCTTGGATGTGGTTCCTATTTCTTCTCTCC	812
	1..2045	Qy	283	ThrArgTyrPheSerIysArgProArgPhePheTyrValSerAlaMetAlaProLeu	302
	/gene="ATST1"	Db	813	ACCAGATATTTTCAGCATCAAGAACCAAAATCTTTTGGTGGCGGATGGCTCTCTTG	872
	33..2009	Qy	303	ThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVal	322
	/codon_start=1	Db	873	ACCTCAGTGTCTTGGAGTCTCTTGGTTTACTTCACTCAGCTCAGAGACATGGTGT	932
	/product="sulfate transporter"	Qy	323	GluValIleGlyLeuLysLysGlyLeuAsnProSerSerLeuThrLeuValPhe	342
	/protein_id="BA21657.1"	Db	933	CAAGTCATAGGGGACCTGAAGAAAGGTTGAATCCACTCTCCGTTCTGATCTCATCTTT	992
	/db_xref="GI:2285885"	Qy	343	ValSerProTyrMetThrAlaValLysThrGlyValValGlyIleLeuSerLeu	362
	/translations="MGTEYTFPQGAELHRRHHTVEAPOQPFLLKSLQSVKELTFP DDPFPQKNQASRKQVGLKFLPPEWAPRYNLKPFKSDLIAGITIIASLAIPQGIS YAKLANLPPIILGYSSPVPPLVAVLGGSSRDLAAGTVAVASLLTGAMLSKEVDAEKDP KLVLHAFATTFPAGVLEASLGI:FRLGFIYDFLSHATIVGPMGAATVVSLOQLKGI GLGHFDTSDIVSMRSVFSQTHREWESVLCGGLFFLLSTRYSIKPKPFVVA MAPLSVILGLSVFYFAERHGVQVIGDLKGLNPLSGSDLIITSPYMSAVTKTGLI TGIIAAGAVGSRFAMFNKYNIDGNKEMIAFGMNVIGSFSCYLTITGPPFSRNV KVKDFPLCMASAVGVVFGSVGLVVAISIALRLFPVRPKTAVKGNIPNSMI RNTQPSRSRTVGILIEDAIFYANASYLRERIRIRWIDEEERVKQSGESSLOYI ILDMASVGNIDTSGISVMVEIKKVIDRRALKVLISNPKGEVVKKLTRSPFIDHLCRE WMFLTVGEAVEACSYMLHTFKTEPASKNEPNNV"	Db	993	ACTTCCCTTACATGTCCACAGCTGTCAAAACTGGGCTCATCACTGGCATCATTGTCTC	1052
Alignment Scores:	1..48-201	Qy	363	AlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGly	382
Pred. No.:	2642.50	Db	1053	GCTGAAGGAATACAGCTGGGGAGAGTTCGCGATGTTCAGAACTACAACATAGACGG	1112
Score:	88.16%	Qy	383	AsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr	402
Percent Similarity:	Conservative: 80	Db	1113	AACAAAGAGATGATAGCGTTTGAATGTGAATGATGAACATCTTCTTTCATCTTGTAC	1172
Best Local Similarity:	Mismatches: 75	Qy	403	LeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAla	422
Query Match:	Indels: 3	Db	1173	CTCACAAACGGACATTTTCAAGTCGGCAGTGAATACACCGGGTTCGAACCGCA	1232
DB:	Gaps: 3	Qy	423	AlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeu	442
US-10-762-049-18 (1-680) x D89631 (1-2045)		Db	1233	ATGTCCAACATAGTAGTGGGATGGGTATGTTTGCACACTCTCTTCTCACACCGCTT	1292
Qy	24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln	Qy	443	PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu	462
Db	33 ATGGGACCGGAGACTACACATTCCTCAAGGACGGGAGGCTTCACCGCGCCACCAC	Db	1293	TTTCACTACACACCACTCGTCGTCCTCTCTGCCATCATCATATCCGCAATGCTCGGACTC	1352
Qy	44 ---ValGluValProProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu	Qy	463	IleAspTyrGluAlaAlaIleHisPheLysValAspLysPheAspPheValValCys	482
Db	93 ACGGTGGAGGCTCCACAACTCTCAACCGTCTTGAAGTCACTTCAGTACTCAGTGAAGAA	Qy			
Qy	63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe				
Db	153 ACTCTGTTTCAGACGACCTTTTAGACAAATTTAAGAACCAAAATGATCAAGAAAATTT				
Qy	83 MetLeuGlyLeuGlnPhePhePheProIlePheGluTyrAlaProLysTyrThrPheGln				
Db	213 GTGTAGGCTCAAACTACTTCTCCGATTTTCAATGGGACCAACCGCTACATCTCAAG				
Qy	103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly				

Db	1353	ATTGACTATCAGCTGCATCTCTCGAAGAGTTGACAAAGTTCGACTTCCTCGTCTGC	1412
Qy	483	MetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle	502
Db	1413	ATGAGCCCTACGTTGGGTCGTATTCGCGAGTGTAGAGATTGGACTCGTCTGAGCGGTG	1472
Qy	503	ValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheValLeuGly	522
Db	1473	CGCATATCTATAGCGAGTGTGCTGCTGTTGTCGAGGCCAAAACCTGCGGTGAAGGA	1532
Qy	523	AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVal	542
Db	1533	ACATACCAACAGCATGATCTATAGGAACACTGAGCAGTACCCTCATCAAGAACCGTT	1592
Qy	543	ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu	562
Db	1593	CCTGGTATCTCATCTTGGAGATTGATGCTCCCATCTACTTTCGTATGCGCAGTTACTTG	1652
Qy	563	ArgGluArgIleThrArgTrrIleAspGluGluGluArgIleLysAlaThrGlyGlu	582
Db	1653	CGTCAAGAATCATAGGTGATGATGAAGAGAGAGAGAGTAAACAATCAGGAGAG	1712
Qy	583	ThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly	602
Db	1713	AGCATGTTACAATATATATCTACGATGATGTCAGCTGTGCTGTAATATCGCACACCGGT	1772
Qy	603	IleSerMetLeuGluGluValLysIleThrGluArgArgGluLeuGlnLeuValLeu	622
Db	1773	ATTAGCATGATGCTGGAAATTAAGAAGTCTATTGACGAGAGCGCTTAAGTTGTTATG	1832
Qy	623	ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis	641
Db	1833	TCAATTCCAAAGAGAGAGGTCGTGAAGAAATTAACACAGATCCAAATTCATCGGTGATCAT	1892
Qy	642	LeuGlyLysIleThrIleThrValGluGluValGluGluValGluGluValGluGluVal	661
Db	1893	TTCGGCAAGAGTGGATGTTCTTAACGTTGAGGAGCAGTGGAGCGCTGTAGCTACATG	1952
Qy	662	LeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrrAsnAsnVal	680
Db	1953	CTTCACACGTTTAAACCGAACCG---GCCTCCAAAACGAGCGCTTGGACACAGTA	2006
RESULT 3			
AY099873			
LOCUS	2266 bp	mRNA	linear
DEFINITION	Arabidopsis thaliana sulfate transporter ATST1 (At3g51895) mRNA,		
complete cds.			
AY099873			
VERSION	AY099873.1	GI:20466813	
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis.		
	1 (bases 1 to 2266)		
	Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M., Palm,C.J., Bower,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.		
TITLE	Submitted (24-APR-2002) DNA sequencing and Technology Center,		
JOURNAL	Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA		
COMMENT	e-mail for correspondence: arab@sequence.stanford.edu		

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki M., Narusaka M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Nguyen, M., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C.J., Bower, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M.K., Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.

#### FEATURES

source

1..2266  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="RAF109-95-N16 (R25423)"  
/ecotype="Columbia"  
/note="this clone is in pBluescript"  
1..2266  
/gene="At3g51895"  
84..2060  
/gene="At3g51895"  
/codon\_start=1  
/product="sulfate transporter ATST1"  
/protein\_id="AA020724.1"  
/db\_xref="GI:20466814"  
/translation="MGTEBDYTPQGAELRRHHTVEAPQPOPFLLKSLQYSVKETLFP  
DDPQFKNQASRFVLGKFLPIFEWAPRYNLKFKSDIAGITIASLAPQGIS  
YAKLANLPILGLSYFPVPLVYSGSRDLAVGTVAVALLTGAMLSKEYDAEKDP  
KLYLHLATPATTAGVLEASLGIFRLGTFVDFLSHATVFGMGAAATVSLQQLKGI  
GLKHTDSTDVISVMRSVFSTHEWRWSGLGCGFLPFLSTRTVFSIKKPFVFAA  
MAPLSTVLGSLVYTHAARGHGOVIGDLKKGLNPLSGDLIFTSPYMSTAVKTCGLI  
TCIIALAREGAVGRSFAMFKYNNIDGNKEMIAFGMMNIVGTSFCTCYLTGPPSRNAV  
YNAGCTAMSNIVMAIAVMFLTLPFLFHYTLPVLVLSAIIISMLGLIDYQAALHLW  
KYDFDFLVCSAYVGVVGSVEIGLVAVASISARLLFVSRPTAVKNGTIPNSMIY  
RNTGEYPSRTVPGILILEIDAPIYFANASYLRERIIRWIDEEERVKSGSSSQYI  
ILDMSAVGNIDTSGISMMVEIKKVDRAALKLVLNPKGVVVKLTRSKFIQDHLGKE  
WMFLTVGEAVEACSVMLTFTKTEPASKNEPNNV"

#### ORIGIN

##### Alignment Scores:

Pred. No.: 1,92e-201 Length: 2266  
Score: 2641.50 Matches: 500  
Percent Similarity: 88.16% Conservative: 81  
Best Local Similarity: 75.87% Mismatches: 75  
Query Match: 76.01% Indels: 3  
DB: 8 Gaps: 3

US-10-762-049-18 (1-680) x AY099873 (1-2266)

Qy	24	MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln	43
Db	84	ATGGGACGCGAGGACTACACATTCCTCAAGAGCGGAGGAGCTTCACCGCGCCACAC	143
Qy	44	---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu	62
Db	144	ACGGTGGAGGCTCCACACCTCAACCGTCTCTTGAAGTCACTTCAGTACTCAGTGAAGNA	203
Qy	63	ThrPhePheProAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe	82
Db	204	ACTCTGTTTCCAGACGACCTTTTAGACAAATTAAGAACCAAAATGCATCAAGAAATTT	263
Qy	83	MetLeuGlyLeuGlnPhePhePheProIlePheGluTrrAlaProLysTyrThrPheGln	102
Db	264	GTGTTAGGCGCTCAATACTTCTCCCGATTTTCGAATGGGACCGCTACATCTCAAG	323
Qy	103	PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly	122
Db	324	TTCCTCAATCAGATCTCATCGCGGAGATCACCATCGCTAGTCTGCCATCTCCTCAGGC	383
Qy	123	IleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIle	142

```
Db 384 ATCAAGTTACGCCAACTTGTAACTTGCCTCCCAATCTTGGCCCTTTATTCAGATTITGTA 443
Qy 143 ProProLeuLeuIleValAlaMetMetGlySerSerArgAspLeuAlaValAlaGlyThrValAla 162
Db 444 CCGCCATTGGTATACCGCGGTCTAGGAGGTTCAAGGACATTAGCGGTGGCAACGGTTGGC 503
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAlaAspProAsnGluAAspPro 182
Db 504 GTTGGCGTCTCTGTGTGACAGGTGGATGCTGAGCAAGAGAGTGTGATGCTGAGAAAGATCCT 563
Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 564 AAGCTTTACCTTACCTTGTCTTTCACCGCCACTTTTTCGCGCGGTCTTCGAGGCTCT 623
Qy 203 LeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleLeGly 222
Db 624 CTTGGAAATTTCAAGTTAGGTTTCATAGTCGACTTCTTATCGCATGCAACGATAGTAGGA 683
Qy 223 PheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeu 242
Db 684 TTCATGGGAGGAGCAGCGATCGGTGGTGGTCTGCAACAGCTTAAGGGTATTTTCGGACTT 743
Qy 243 GluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThr 262
Db 744 AACATTTTCAGACTCTACGATGTTATCTGTGATGCTTCCGTTTCTCCCAACT 803
Qy 263 HisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeuLeuSer 282
Db 804 CACGAGTGGAGATGGGAAAGTGGGTTCTTGGATGGTCTTATTTCTTCTTCTCTCC 863
Qy 283 ThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLeu 302
Db 864 ACCAGATATTTCAAGATCAAGAACCAAAATCTTTTGGGTGGCGGATGGCTCTCTTG 923
Qy 303 ThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVal 322
Db 924 ACCTCAGTGAATCTTGAAGATCTCTTGGTTTACTTCACTCACGCTGAGACATGGTGT 983
Qy 323 GluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPhe 342
Db 984 CAAGTATAGGGACCTGAGAAAGGTTGAATCCACTCTCCGGTCTGATCTCATCTTT 1043
Qy 343 ValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleLeuSerLeu 362
Db 1044 ACTTCCCTTACATGTCCACAGCTGTCAAACTGGCTCATCACTGGCATCAITGTCTC 1103
Qy 363 AlaGluGlyIleAlaValAlaGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGly 382
Db 1104 GCTGAAGGATAGCAGTGGGAGAGGTTTGGCATGTTCAAGAACTACAAACATAGACGGG 1163
Qy 383 AsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyr 402
Db 1164 AACAAAGATAGATAGGTTGGATGATGAAATCATCGTTGGTTCCTTCATCTTGTATC 1223
Qy 403 LeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAla 422
Db 1224 CTCACACCGGACCATTTTCAAGTCGCGAGTGAATACAAACGCGGTTGCAAGACGCA 1283
Qy 423 AlaSerAsnIleLeuMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeu 442
Db 1284 ATGTCCAAACATAGTATGGCGATTTGGGTTATGTTTCACATCCTCTTCTTCACACCGCTT 1343
Qy 443 PheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeu 462
Db 1344 TTTCACTACACACCATCTGTCGCTCTCTGCGCATCATCATATCCGAAATGCTCGACTC 1403
Qy 463 IleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCys 482
Db 1404 ATTGACTATCAAGCTGCCATCTCTCGAAAGTTGCAAGTTTCGACTTCTCTCGTCTGC 1463
Qy 483 MetSerLalaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIle 502
```

```
Db 1464 ATGAGCGCTACCTGGGTGCTATTCCGCGAGTGTAGATTGGACTCGTCTGAGCGGTG 1523
Qy 503 ValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGly 522
Db 1524 GCGATATCTATGAGGAGTGTGTGCTTGTTCGAGGCCAAATACTGCGGTGAAGGA 1583
Qy 523 AsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaHisVal 542
Db 1584 AACATACCAAAACAGCATGATCTATAGAACCTGAGCAGTACCCTCATCAAGAACCGTT 1643
Qy 543 ProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeu 562
Db 1644 CTTGGTATTCTCATCTTGGAGATTGATGCTCCCATCTTCTTCTTAATGCCAGTACTTG 1703
Qy 563 ArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGlu 582
Db 1704 CGTGAAGAAGATCATAGGTGGATTTGATGAAGAGAGAGAGAGATTAAACAATCAGGAGAG 1763
Qy 583 ThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGly 602
Db 1764 AGCAGCTTACAATATATATCTGATATGTCAGCTGTGTTGTAATATCGACACACGCGT 1823
Qy 603 IleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeu 622
Db 1824 ATTAGCATGATGCTGGAATTAAGAAGCTCATTCACAGGAGCGCTTAAGTTGGTATTG 1883
Qy 623 ValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHis 641
Db 1884 TCAATCCAAAAGGAGAGGTCGTGAAGAATAAACAGATCCAAATTCATCGGTGATCAT 1943
Qy 642 LeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661
Db 1944 TTGGGCAAGAGTGGATGTTCTTAACGCTAGGAGAACGATGGAGGCTTGTACTCATG 2003
Qy 662 LeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTrpAsnAsnVal 680
Db 2004 CTTCAACGTTTAAACCGAACCG---GCCTCCAAAACGAGCCTTGGAAACAACGTA 2057

RESULT 4
BNA581745 2193 bp mRNA linear PLN 02-SEP-2003
LOCUS Brassica napus mRNA for sulphate transporter (bst3.1 gene).
DEFINITION AJ581745
ACCESSION AJ581745.1 GI:34481597
VERSION bst3.1 gene; sulphate transporter.
KEYWORDS Brassica napus (rape)
SOURCE Brassica napus
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1
AUTHORS Buchner,P., Stuiver,E., Hawkesford,M.J. and de Kok,L.J.
TITLE Analysis of the effect of sulphur nutrition and atmospheric H2S
exposure on the expression of sulphate transporter genes in Curley
Cale
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2193)
AUTHORS Buchner,P.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-2003) Buchner P., Agricultural and Environmental
Division, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ,
UNITED KINGDOM
FEATURES
source
1..2193
/organism="Brassica napus"
/mol_type="mRNA"
/variety="Starlight"
/db_xref="taxon:3708"
1..2193
/genes="bst3.1"
1..79
/genes="bst3.1"
80..2056
CDS
gene
5'UTR
CDS
```





```

Db 1727 GAAGAGAGATAGGATCAAGAGCATCAGGAGGAATAGTTTACATATGTTATCTCGAT 1786
Qy 592 MetSerAlaValGlyAenIleAspThrSerGlyIleSerMetLeuGluValLysLys 611
Db 1787 ATGCTGCTGTTGTAATATCGACACGAGTGGTATAGCATGATGAGGAAATTAAGAA 1846
Qy 612 IleThrGluArgGluLeuGlnLeuValLeuValAsnProValSerGluValMetLys 631
Db 1847 ATCATGGACAGAGAGAGTTTAAAGTTAGTATTGGCAATCCAAAGGAGAGGTCGTGAAG 1906
Qy 632 LysLeuAsnLysSerLysPhe---GlnAsnHisLeuGlyLysLysTrpIleTyrLeuThr 650
Db 1907 AACTTACACAGATCCAAATTCATCGGTGAAATTTGGCCAAAGATGGATGTTCTTAACA 1966
Qy 651 ValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerLysThrAsnProLys 670
Db 1967 GTTGGAGAGCCGTGGAGGCTTGTAGTTTCATGCTTCACACGTCAAAGACCGAACCGGCC 2026
Qy 671 LysAspGluThrGluGlyTrpAsnAsnVal 680
Db 2027 TCCAAAGAA---GAGCCTTGGAAACAGTT 2053

RESULT 5
AX506807
LOCUS AX506807 1977 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1502 from Patent W00216655.
ACCESSION AX506807
VERSION AX506807.1 GI:23388044
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 1502 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source 1. 1977
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
Pred. No.: 7,67e-198 Length: 1977
Score: 2595.50 Matches: 498
Percent Similarity: 87.73% Conservative: 81
Best Local Similarity: 75.45% Mismatches: 77
Query Match: 74.69% Indels: 5
DB: 6 Gaps: 3

US-10-762-049-18 (1-680) x AX506807 (1-1977)
Qy 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43
Db 1 ATGGGCAGGAGGACTACACATTTCCCTCAAGGAGCGGAGAGCTTCACCGCCGCCACCAC 60
Qy 44 ---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db 61 ACGGTGGAGGCTCCACACCTCAACCGTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 120
Qy 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAsnLysLysPhe 82
Db 121 ACTCTGTTTCCAGACGACCCCTTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT 180
Qy 83 MetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPheGln 102
Db 181 GTGTTAGGCGCTCAATACTTCTCCCGATTTTCGAATGGGCGCACCGCTACAAATCTCAAG 240

```

```

Qy 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db 241 TTCTTCAAATCAGATCTCATCGCGGAATCACCATCGTAGTCTGCCATCCCTCAGGCG 300
Qy 123 IleSerTyrAlaIleValLeuAlaAsnLeuProIleLeuLeuGlyLysTyrSerSerPheIle 142
Db 301 ATCAGTTACGCCAAACCTTGTAACTTGGCCCCCAATTTCTGGCCCTTTATTTCGAGTTTGT 360
Qy 143 ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla 162
Db 361 CGCCATTGGTATACCGCGGTAGGAGTTCAAGGAGCTTAGCGGTGGGAACGGTTGCG 420
Qy 163 ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro 182
Db 421 GTTGCCTCTCTGTTCACAGTGCATGCTCAGCAAGAAAGATTGATGCTCAGAAAGATCTCT 480
Qy 183 LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla 202
Db 481 AAGCTTTACCTTCAACCTTGTTCACCGCCACTTTTTCGCCGCGCTTCTCGAAGCCTCT 540
Qy 203 LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGln 222
Db 541 CTGGGAATTTTCAAGGTTAGGTTTCATAGTGGACTTTCTATCGCATGCAACGATAGTAG 600
Qy 222 YPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLe 242
Db 601 ATTCATGGGAGGAGCAGCGCGTGGTGTCTGCAACAGCTTAAGGGTATTTTCGACT 660
Qy 242 uGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnTh 262
Db 661 TAAACATTTTCACAGACTCTACCGATGTATCTGTGTATGTCGTTCCGTTTCTCCCAAC 720
Qy 262 rHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe 282
Db 721 TACGAGTGGAGATGGGAAGTGGCGTCTTGGATGGTCTTCTATTCTTCTTCTC 780
Qy 282 rThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaProLe 302
Db 781 CACCAGATATTGAGCATCAAGAAACCAAAATTTCTTGGTGGCGCGCATGCTCCTTT 840
Qy 302 uThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyVa 322
Db 841 GACCTCAGTGATCTTGGAAAGTCTCTTGTGTTACTTCTACTCAGCTGAGAGACATGGTGT 900
Qy 322 lGluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPh 342
Db 901 TCAAGTGATAGGGACCTGAAGAAAGGTTGAATCCACTCTCCGGTTCTGATCTCATCTT 960
Qy 342 eValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleSerLe 362
Db 961 TACTTCCCTTACATGTCACAGCTGTCAAAACTGGCCCTCATCTCCTGGCATCATTTGCTCT 1020
Qy 362 uAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGln 382
Db 1021 CGC-GAAGGAGTAGCAGTGGGAGGAGTTTTCGATGTTTCAAGAACTCAACACATACACGG 1079
Qy 382 YAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTy 402
Db 1080 GAACAAAGAGATAGTAGCGTTTGGAAATGATGAACATCGTTGGTTCCTTTCATCTTGTTA 1139
Qy 402 rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAl 422
Db 1140 CCTCAACACCGGACCATTTTCAAGTTCGCGAGTGAACACTCAACCGGGTTGCAAGACCGC 1199
Qy 422 AlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheIleuThrProLe 442
Db 1200 AATGTCCAACTAGTAGCGGATTTGCGGTTATGTTTACACTCTCTCTTCTCACACCGCT 1259
Qy 442 uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe 462
Db 1260 TTTTCTACTACACCACTCGTCTCTCTCTGCGATCATCATATCCGCAATGCTCGGACT 1319

```

Qy	462	uileAspTyrGluAlaAlaIleHisLeuPheHisValAspLysPheAspPheValValCy	482
Db	1320	CAATTGACTATCAAGCTGCCATCCATCTCTGGAAGTTGACAACTTCGACTTCCTCGTCTG	1379
Qy	482	sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI	502
Db	1380	CATGAGCGCTACGTTGGGTGCTATTTCGCGACTGTAGAGATTGGACTCGTCTAGCGGT	1439
Qy	502	eValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGI	522
Db	1440	GGCGATATCTATAGCGAGGTGTGTGCTTTGTGTCGAGGCCAAACATCGCGGTGAGGG	1499
Qy	522	yAenIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa	542
Db	1500	AAACATACCAACAGCATGATCTATAGGAACACTGAGCAGTACCCCATCATCAAGAACCGT	1559
Qy	542	lProGlyMetLeuIleGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrIle	562
Db	1560	TCTGTGTATCTCATCTTGAGATTGATGCTCCCATCTACTTTGCTTAATGCCAGTTACTT	1619
Qy	562	uArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThrGlyGI	582
Db	1620	GCCTGAAGAATCATAGGTGGATTGATGAGAGGAAGAGAGATTAAACAATCAGGAGA	1679
Qy	582	uThrSerLeuGlnTyrValIleIleAspMetSerAlaValAlGlyAsnIleAspThrSerGI	602
Db	1680	GAGCAGCTTACAAATATATATATCTGATATGTCAGCTGTGTGGTAAATCGACACAAACGG	1739
Qy	602	yIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValle	622
Db	1740	TATTAGCATGATGGTGAATTAAGAAAGTCATTGACAGAGAGCGGTAAAGTTGTGATT	1799
Qy	622	uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHi	641
Db	1800	GTCAATCCAAAGAGAGAGGTCTGAGAAATTAACAGATCCAAATTCATCGTGATCA	1859
Qy	641	sLeuGlyLysLysTyrPileTyrLeuThrValGluGluAlaValAlGlyAlaCysAsnPheAs	661
Db	1860	TTTGGGCAAGAGTGGATGTTCTTAACGGTAGGAGAAGCAGTGGAGGCTGTAGCTACAT	1919
Qy	661	nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTyrAsnAsnVal	680
Db	1920	GCTTCACAGCTTTAAACCGAACCG---GCCTCCAAACAGAGCCTTGGAAACAACGTA	1974
RESULT	6		
LOCUS	AX651748	1977 bp	DNA
DEFINITION	Sequence 589 from Patent WO0300898.		linear
ACCESSION	AX651748		PAT 22-MAR-2003
VERSION	AX651748.1	GI:29154566	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
TITLE	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
JOURNAL	Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,		
FEATURES	Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.		
source	Plant genes involved in defense against pathogens		
	Patent: WO 0300898-A 589 03-JAN-2003;		
	Syngenta Participations AG (CH)		
	Location/Qualifiers		
	1. .1977		
	/organism="Arabidopsis thaliana"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3702"		
ORIGIN			
Alignment Scores:			
Pred. No.:	7, 67e-198	Length:	1977
Score:	2595.50	Matches:	498
Percent Similarity:	87.73%	Conservative:	81

Best Local Similarity:	75.45%	Mismatches:	77
Query Match:	74.69%	Indels:	5
DB:	6	Gaps:	3
US-10-762-049-18 (1-680) x AX651748 (1-1977)			
Qy	24	MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln	43
Db	1	ATGGGCACGGAGGACTACACATTCCTCAAGGAGCGGAGGAGCTTCACCGCGCCACCAC	60
Qy	44	---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu	62
Db	61	ACGGTGGAGGCTCCCAACCTCAACCGTTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA	120
Qy	63	ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe	82
Db	121	ATCTGTTTCCAGACGACCCCTTTAGACAATTTAGAACCAAAATGCATCAGAAAATTT	180
Qy	83	MetLeuGlyLeuGlnPhePheProIlePheGluTyrAlaProLysTyrThrPheGln	102
Db	181	GTGTTAGGCTCAAAATACTTCTCCGATTTTCGAATGGGCACCGCTACAATCTCAAG	240
Qy	103	PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly	122
Db	241	TTCTTCAAAATCAGATCTCATCCGCGAATCACCATCGCTAGTCTCGCCATCCCTTCAGGC	300
Qy	123	IleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPheIle	142
Db	301	ATCAGTTACGCCAAACTTGTCTAACTTGGCCCCCAATCTTGGCTTTATTTCGAGTTTGT	360
Qy	143	ProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAla	162
Db	361	CGCCATTGGTATACGGGTGCTAGGAGTTCAAGGAGCTTAGCGGTGGGAACGGTTGCG	420
Qy	163	ValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspPro	182
Db	421	GTTCGCTCTCTGTGACAGGTGCGATGCTGACCAAGAGTTGATGCTCAGAAAGATCT	480
Qy	183	LysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAla	202
Db	481	ANGCTTTACCTTCACCTTGTCTTCCCGCCACTTTTTCGCCGGGGTTCGAAAGCTCT	540
Qy	203	LeuGlyLeuPheArg-LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGI	222
Db	541	CTTGGAAATTTTCAAGGTTAGGTTTATAGTGACTTTCTATCGCATGCAACGATAGG	600
Qy	222	yPheMetGlyGlyAlaAlaThrValValCysLeuGlnLeuLysSerIleLeuGlyLe	242
Db	601	ATTTCATGGGAGGAGCAGCGAGTGTGAGTCTGCAACAGCTTAAAGGGTATTTTCGGACT	660
Qy	242	uGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnTh	262
Db	661	TAAACATTTACAGACTCTACCGATGTTATCTCTGTGATCGTTCGTTTTCTCCCAAC	720
Qy	262	rHisGluTyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSe	282
Db	721	TCACGAGTGGAGATGGGAAAGTGGGTTCTTGATGTGGTTTCTTATTCTTCTCTCTC	780
Qy	282	rThrArgTyrPheSerLysLysArgProArgPhePheThrValSerAlaMetAlaProle	302
Db	781	CACCAAGATTTTGGAGCATCAAGAAACCAAAATCTTTTGGGTGGCGGATGGCTCCTTT	840
Qy	302	uThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGlyVa	322
Db	841	GACCTCAGTATCTTGGAGTCTCTTGTGTACTTCTACTCAGCTGAGAGACATGGTGT	900
Qy	322	IgluValIleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPh	342
Db	901	TCAAGTGATAGGGACCTGAAGAAAGGGTTGAATCCAATCTCTCCGGTTCTGATCTCATCT	960
Qy	342	eValSerProTyrMetThrAlaValLysThrGlyIleValValGlyIleIleSerIe	362
Db	961	TACTTCCCTTACATGTCCACAGCTGTCAAAACTGGCCTCATCATCTGGCATCATGTCT	1020

Qy	362	uAlaGluGlylleAlaValAlGlyArgSerPheAlaMetTyrIysAsnTyrAsnIleAspG1	382
Db	1021	CGC-GAAGSAGTA3CAGTGGGAGGAGTTTTGGCATGTTCAAGAACTACAACTAGACGG	1079
Qy	382	YAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTy	402
Db	1080	GAACAAAGAGATGATACGTTTGGATGATGAACATCGTTGGTCTCTTCACTCTTGTTA	1139
Qy	402	rLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysIysThrAl	422
Db	1140	CCTCACAAACCGACATTTTCAAGTCCGCGAGTGAACACTCAACGCGGGTTGCAAGACCG	1199
Qy	422	aAlaSerAsnIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLe	442
Db	1200	AATGTCCAACATAGTAGCGATGGCGTTATGTTTCACACTCCTCTTCTCCACCGCT	1259
Qy	442	uPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLe	462
Db	1260	TTTTCACTACACCACTCGTCTCTCTGCGATCATCATATCCGCAATGTCGGACT	1319
Qy	462	uIleAspTyrGluAlaIleHisLeuPheLysValAspLysPheAspPheValValCy	482
Db	1320	CATTGACTATCAAGTCCCATCTCTCGAAAGTTGACAAAGTTCGACTTCTCTCGTCTG	1379
Qy	482	sMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaI	502
Db	1380	CATGAGCGCTACGTTGGGGTCTGTTTCCGACGTGTAGAGATTGGACTCGTCTAGCGGT	1439
Qy	502	eValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeuG1	522
Db	1440	GGCGATATCTATACGAGGTGTGCTGTTTGTGCGAGCCAAACCTCCGGTGAAGGG	1499
Qy	522	YAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisVa	542
Db	1500	AAACATACCAACACGACATGATATAGGAACACTGACGAGTACCCTCATCAAGAACCGT	1559
Qy	542	lProGlyMetLeuLeuGluLeuAspAlaProIleTyrPheAlaAsnAlaSerTyrLe	562
Db	1560	TCCTGGTATTCTCATCTTGGAGATTGATGCTCCCATCTACTTGTCTAATGCCAGTTACTT	1619
Qy	562	uArgGluArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyG1	582
Db	1620	GCGTGAAGAAGTATATAAGTGGATGATGAAGAGGAGAGAGAGTTAAACAATCAGGAGA	1679
Qy	582	uThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerG1	602
Db	1680	GACGAGCTTACATATATTACTCGATATGTCAGCTGTTGTTGTTAATATCGACACAGCGG	1739
Qy	602	YIleSerMetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLe	622
Db	1740	TATTAGCATGATGTTGAAATTAAGAAAGTCATTGACAGGAGAGCGTTAAAGTTGGTATT	1799
Qy	622	uValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPhe---GlnAsnHi	641
Db	1800	GTCAATATCCAAAGAGAGAGTCTGGAAGAAATTAACCGATGCCAATTTATCGGTGATCA	1859
Qy	641	sLeuGlyLysLysTrpIleTyrLeuThrValGluGluAlaValAlGlyAlaCysAsnPheAs	661
Db	1860	TTTGGGCANAGATGGATGTTCTTAACGGTAGGAGAGCAGTGGAGGCTTGTAGCTACAT	1919
Qy	661	nLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTyrTrpAsnVal	680
Db	1920	GCTTTCACACGCTTTAAACCCGAACCG---GCCTCCAAAACGAGCGCTTGAACAACGTA	1974
RESULT 7			
AX653455			
LOCUS	AX653455	1974 bp	DNA
DEFINITION	Sequence 3325 from Patent W003000898.	linear	PAT 22-MAR-2003
ACCESSION	AX653455		
VERSION	AX653455.1	GI:29156269	
KEYWORDS	Oryza sativa		
SOURCE			

ORGANISM		Oryza sativa	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.	
AUTHORS		Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M., Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.	
TITLE		Plant genes involved in defense against pathogens	
JOURNAL		Patent: WO 0300898-A 3325 03-JAN-2003;	
FEATURES		Syngenta Participations AG (CH)	
source		Location/Qualifiers	
		1..1974	
		/organism="Oryza sativa"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:4530"	
ORIGIN			
Alignment Scores:		1..176-176	
Pred. No.:		Length:	
Score:		2330.50	
Percent Similarity:		Matches:	
Best Local Similarity:		Conservative:	
Query Match:		Mismatch:	
DB:		Indels:	
		Gaps:	
US-10-762-049-18 (1-680) x AX653455 (1-1974)			
Qy	43	GlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu	62
Db	52	AGGTCGCCGATGCGCGCGGAGCGTCTCTCGAGACGCTGGGGGGAACATGAAGGAG	111
Qy	63	ThrPhePheProAspAspProLeuArgGlnPhe---LysAsnLysProAlaSerLysLys	81
Db	112	ACATTTCTCGCGACACCGCTTCAAGGTGTGTCGCGGAGCGCGGTGTCGCGGCGGC	171
Qy	82	PheMetLeuGlyLeuGlnPhePheProIlePheGluTrpAlaProLysTyrThrPhe	101
Db	172	GCGCGCGCGCGCTCCGGTACGTGTTCCCGTTCATGGAGTGGCGCGCTCGTACACCTC	231
Qy	102	GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln	121
Db	232	GGCACCTCAAGTCCGACCTCATCGCGGACATCACCATTGCCAGCTCGCCATCCCCAG	291
Qy	122	GlyIleSerTyrAlaLysLeuAlaAsnLeuProIleLeuGlyLeuTyrSerSerPhe	141
Db	292	GGCATCAGCTAGCCCAAGCTCGCAACCTCCCTCCGCTCTCGGCTCTATTTCGAGTTC	351
Qy	142	IleProProIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal	161
Db	352	GTGCCCGCGCTGGTGTACGCGATGATGGGAGCTCGAGGACCTGGCGGTGGGAGCGGTG	411
Qy	162	AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp	181
Db	412	GCGGTGCGCTCGCTGCTGATCGGTGCTGATGAGGAGGAGGTGTCGCGCGGAGGAC	471
Qy	182	ProLysLeuTyrIleHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla	201
Db	472	CGGCGCTGTACCTGACGCTCGCGCTCACCGCCACCTTCTTCCGCGCGGTTCACAGCG	531
Qy	202	AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle	221
Db	532	CTGCTGGGCGCTTGGAGCTGGGCTTCATCGTGGACTTCTGTGCGACCGCCATCGTC	591
Qy	222	GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly	241
Db	592	GGGTTTCATGGCGCGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	651
Qy	242	LeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln	261
Db	652	CTCGACCACTTCAACACCGCCACCGACCTCGTCTCGCTCATGAGCTCCGCTCTTCCCA	711
Qy	262	ThrHisGluTrpArgTrpGlnSerAlaValLeuGlyCysValPheIlePheLeuLeu	281

Db 712 ACCACCTCTGGCGATGGGAGACGCTGTCATANGGCTGGCGCTTCCTCTTCCTCCCTC 771  
Qy 282 SerThrArgTyrPheSerLysLysArgProArgPhePheTyrValSerAlaMetAlaPro 301  
Db 772 ATCACCGCTTCTTCAGCAAGAGGAGCCNAGGTTCTTCTGGGTATCTGCAGCTGGCCA 831  
Qy 302 LeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGluLysHisGly 321  
Db 832 TTGGCGTCTGTATCATCGGAGGCTCTGTTGTACTCTCACATCATGCTGAAACCATGCG 891  
Qy 322 ValGluValIleGlyGluLeuLysLysGlyLeuAsnProSerLeuThrAsnLeuVal 341  
Db 892 ATTCAGATGATTGTTACCTGGAAGAGGCGCTGAACCCACCGCTCTCCCAAGCCCTGAAC 951  
Qy 342 PheValSerProTyrMetThrThrAlaValLysThrGlyLeuValValGlyLysLeuSer 361  
Db 952 TTCTCTGCGCGTACATGATGCTGGCCCTGAAGACCGGATCATCTACTGCGGTATGTC 1011  
Qy 362 LeuAlaGluGlyLeuAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAsp 381  
Db 1012 CTGCTGAGGGATTGCGGTAGGAAGGAGCTTTGGCAATGTTCAAGAAATTACACATCGAC 1071  
Qy 382 GlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCys 401  
Db 1072 GGAACAAGAGAGATGATCGGTTTGAACGATGAACATTGTGGGATCATCTACCTCTTGC 1131  
Qy 402 TyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThr 421  
Db 1132 TACCTCACACCGTCTGTTCTCGAGTGGCGCTCAATCATCAACCGCGCTGCAGAGC 1191  
Qy 422 AlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrPro 441  
Db 1192 GCGATGCGAAACGTGATCATGTCGTCGCGGTGATGATCAGCTGCTGTTCTTGACGCG 1251  
Qy 442 LeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGly 461  
Db 1252 CTGTTCCACTACACCGCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1311  
Qy 462 LeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValVal 481  
Db 1312 CTGATCGACTACCGCGCGCGCTCCACTGTGGCAGGTGGACNAGGTGCATCTTCGCTC 1371  
Qy 482 CysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAla 501  
Db 1372 TGCTCGCGCGCTTACCTCGCGCTGCTTTCGCGCGCTCGAGATCGCTCGTCGCGC 1431  
Qy 502 IleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeu 521  
Db 1432 GTCGGATCTCCATCTCCGCGTCTGCTGTTGTCGCGCGCGCGAGACGAGTGTCTC 1491  
Qy 522 GlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHis 541  
Db 1492 GGGAACTCCCNACAGATGATCTACCGCGGATGGACGATACCCCGCGGCGAGG 1551  
Qy 542 ValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyr 561  
Db 1552 GTCCCCGCGTGTCTGCTCCGCGTCGATCGCTCCCATCTACTTCCACCAACGCCAGCTAC 1611  
Qy 562 LeuArgGluArgIleThrArgThrIleAspGluGluGluGluArgIleLysAlaThrGly 581  
Db 1612 CTGCGTGAAGGATCCGCGGTGGATCGACGAGGAGGACCATCGTCGAAGGAGGCGC 1671  
Qy 582 GluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSer 601  
Db 1672 GAGATGGGCATTCAGTACGTCTCTCGATGGGTGGGTGGAGGATCCGATCCACCCAGC 1731  
Qy 602 GlyIleSerMetLeuGluGluLysLysIleThrGluArgArgGluLeuGlnLeuVal 621  
Db 1732 GGGACGAGCATGCTGGATGAATCAGCAAGACCCCTGGACAGGAGGCGCTTCAGATTGTG 1791  
Qy 622 LeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHis 641  
Db 1792 TTGGCGAACCCCGGGGAGGAGATCATGAAGAGTTGGACAGCTCCAGGTGCTTGAGGCG 1851

Qy 642 LeuGlyLysLysTyrIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsn 661  
Db 1852 ATCGGCCATGATGATGATCTTCCCCAGGTGGCGAGCGGTGGCGACTTCGTCG 1911  
Qy 662 LeuArgAlaSerLys 666  
Db 1912 ATGCACCTCGCAGAAG 1926  
RESULT 8  
AK066932  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) cDNA clone:J013093F02, full insert sequence.  
ACCESSION  
AK066932  
VERSION  
AK066932.1 GI:32976950  
KEYWORDS  
FLJ\_CDNA; CAP trapper.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE  
1  
The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shisniki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtsuka, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurotsaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Koyama, S., Kurihara, C., Kurotsaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namioka, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ohtsuka, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.  
Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kanondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)  
TITLE  
JOURNAL  
MEDLINE  
22752273  
PUBMED  
12869764  
REFERENCE  
2 (bases 1 to 2327)  
AUTHORS  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurihara, C., Kurotsaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namioka, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ohtsuka, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.  
TITLE  
JOURNAL

COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice. URL : <a href="http://cdna01.dna.affrc.go.jp/cDNA/">http://cdna01.dna.affrc.go.jp/cDNA/</a> NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Iehikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M. FAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Iehibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Nariikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Teunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Oeato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y. Location/Qualifiers	
FEATURES	1..2327 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /clone="J013093P02"	
source		
ORIGIN		
Alignment Scores:		
Pred. No.:	1.44e-176	Length: 2327
Score:	2330.50	Matches: 441
Percent Similarity:	83.84%	Conservative: 83
Best Local Similarity:	70.56%	Mismatches: 100
Query Match:	67.06%	Indels: 1
DB:	8	Gaps: 1
US-10-762-049-18 (1-680) x AK066932 (1-2327)		
Qy	43	GlnValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62
Db	163	AGGGTGGCGATGCGCGCGGCGAAGCCGTTCTGGAGACGCTGGGGGGGACATGAAGGAG 222
Qy	63	ThrPhePheProAspAspProLeuArgGlnPhe---LysAsnLysProAlaSerLysLys 81
Db	223	ACATTCTGCGGACGACCGCTTCAGGGTGGTGGCGGGGAGCGCGGGTGGCGGGCGG 282
Qy	82	PheMetLeuGlyLeuGlnPhePhePheProIlePheGluTrpAlaProLysTyrThrPhe 101
Db	283	CGCGCGCGCGGCTCCGGTACGTTCCCGTTTCATGGAGTGGCGCGCGTGTACACCCCTC 342
Qy	102	GlnPheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGln 121
Db	343	GGCACCCCTCAAGTCCGACCTCATCGCGCGCATCACCATTGCCAGCCTCGCCATCCCCCAG 402
Qy	122	GlyIleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeuTyrSerSerPhe 141
Db	403	GGCATACGACTACGCCAAGCTCGCCACACTCTCTCTCCCGTCTCTGGCCTCTATTGAGCTTC 462
Qy	142	IleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVal 161
Db	463	GTGCGCGCGCTGTGNTACGCGATGATGGGAGCTCGAGGGACCTGGCGGTGGGACGGTG 522
Qy	162	AlaValGlySerLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAsp 181
Db		
Qy	523	CGGGTGGCGTCCGTGCTGATCGGGTCGATCGAGCGAGGAGGTGTGGCGGGGAGGAC 582
Qy	182	ProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAla 201
Db	583	CGGGCGCTGTACCTGCACGCTCGCGCTCACCGCCACCTTCTTCCGCCGCGTGTTCAGGCG 642
Qy	202	AlaLeuGlyLeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIle 221
Db	643	CTGCTGGGGGTCTTGAGGCTTGGGGTTCATCGTGGATCTTCTGTCCAGCGCCACATCGTC 702
Qy	222	GlyPheMetGlyGlyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGly 241
Db	703	GGGTTTCATGGCGCGCGCCCGCCCGTGGTGTGCTCGCAGCAGCTCAAGGGCATGTTCCGC 762
Qy	242	LeuGluHisPheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGln 261
Db	763	CTCGACCACTTCCACCACCGCCACCGACCTCGTCCGTCATGAGCTCCGTCTTCTCCCAA 822
Qy	262	ThrHisGluTrpArgTrpGluSerAlaValLeuGlyCysValPheIlePhePheLeuLeu 281
Db	823	ACCACTCTGGCGATGGGAGAGCGTGCATCGGCTGGCGCTTCCTCTTCTTCTCTCTC 882
Qy	282	SerThrArgTyrPheSerLysLysArgProArgPhePheTrpValSerAlaMetAlaPro 301
Db	883	ATCACCGGCTTCTCAGCAAGAGGAGGCCAAGGTTCTTCTGGGTATCTGCAGCTGGGCCA 942
Qy	302	LeuThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysHisGly 321
Db	943	TTGGGCTGTCTCATCATCATCGGAGCCTCTCGGTGTACTCACATGCTGAAACCATGGC 1002
Qy	322	ValGluValIleGlyGluLeuLysGlyLeuAsnProProSerLeuThrAsnLeuVal 341
Db	1003	ATTCAAGTGTATGGTTTACTCGAAGAGGGCGCTGAACCCACCGCTGCCCAACCGCTGAAC 1062
Qy	342	PheValSerProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSer 361
Db	1063	TTCTCTCGCGCTACATGATGCTGGGCCCTGAAGCCGGGATCATCATCTGGCGTCATTGCC 1122
Qy	362	LeuAlaGluGlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAsp 381
Db	1123	CTCGCTGAGGGGATTCCGCTAGGAGGAGCTTTGCAATGTTCAAGAAATTACCACATCGAC 1182
Qy	382	GlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCys 401
Db	1183	GGAAACAAGGAGATGATCGCGTTTGGAAACGATGAACATTTGTGGGATCATCTCACCCTCTTC 1242
Qy	402	TyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThr 421
Db	1243	TACCTCACCAACCGTCCGTTCTCGAGTGGCGCTCACTACCAACCGCGGCTGCAAGACG 1302
Qy	422	AlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrPro 441
Db	1303	CGGATGTCGAACGTGATCATGTCGGTGGCGGTGATGATCATCGCTGCTGTTCTTTCAGCGCG 1362
Qy	442	LeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGly 461
Db	1363	CTGTTCCACTACACGCGCTGGTGTGTGTCTCGCGATCATCATGTGCGCGATGCTGGGG 1422
Qy	462	LeuIleAspTyrGluAlaIleHisLeuPheLysValAspLysPheAspPheValVal 481
Db	1423	CTGATCGACTACCGCGCGCGCTCCACCTGTGGCAGGTGGCAAGGTGGACTTCTTCGGTTC 1482
Qy	482	CysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAla 501
Db	1483	TGGCTCGGCGCTACCTCGCGCTCGTCTTCCGAGCGTTCGAGATCGGCTCGTCTCGTCCGC 1542
Qy	502	IleValIleSerValLeuArgValLeuLeuPheIleAlaArgProArgThrPheValLeu 521
Db	1543	GTGCGGATCTCCATCTCCGCGTGTGTGTCTCGTGGCGCGCGGAGGACGACGCGTCTC 1602
Qy	522	GlyAsnIleProAsnSerValIleTyrArgAsnValGluHisTyrGlnAlaLysHis 541
Db	1603	GGGAACATCCCCAACACAGATGATCTACCGCGCGGATGGACCGATACACCGCGCGGACGAGG 1662







Db	907	GGTATCCAAATTATTCGGAGAACTAAAGAAAGGATAAATCCACCGTCATCACACATTTG	966
Qy	341	ValPheValSerProTyrMetThrAlaValLysThrGlyLeValValGlyLeIle	360
Db	967	GTITTCACGCCACCTATGATGTGGCTCTCAAGTTGGCATATCACTGGAGTCATA	1026
Qy	361	SerLeuAlaGluGlyLeAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIle	380
Db	1027	GCTCTGCTGAGGTATAGCAGTGGGAGAAAGCTTTGCAATGTACAGAAATTAATATA	1086
Qy	381	AspGlyAsnLysGluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSer	400
Db	1087	GATGGGAACAAGAGATAGATGCTTTGGGATGATGAACATCTTGTTCTCTCTCTCT	1146
Qy	401	CysTyrLeuThrThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLys	420
Db	1147	TGCTATCTCTACTACCGGGCCATTTTCACGTTCCGGCGTGAACATACACGCGGGTGC	1206
Qy	421	ThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr	440
Db	1207	ACGGCGTTATCAACGTTGGTGATGGCGTTCGAGTGGCGGTGACGTTCCTTTTAACG	1266
Qy	441	ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu	460
Db	1267	CCGTTATCTTCTACACGCCATTAGTCTGTCGTCGATCATATCGTCGCATGCTA	1326
Qy	461	GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal	480
Db	1327	GGCGTAGTGACTACGAAAGCAGCATTCATCTTTGGAACTCGACAAAGTTCGATTTCTTC	1386
Qy	481	ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle	500
Db	1387	GTITGTCCTCTCTACTTGGCGGTGTTTGGTACCATCGAAATTTGGTCTCATCCTC	1446
Qy	501	AlaIleValIleSerValLeuArgValLeuPheIleAlaArgProArgThrPheVal	520
Db	1447	TCGCTGGGAATACAGTAGAGGTAGTGTCTGTCGGAGAGCAAGATTATGTA	1506
Qy	521	LeuGlyAsnIleProHisSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLys	540
Db	1507	ATGGGAACAATTCAAAACCTCAGAAATATATAGGAACATTCAGCATTCACCCCAAGC	1566
Qy	541	HisValProGlyMetLeuIleLeuGluIleAspAlaProIleTyrPheAlaAsnAlaSer	560
Db	1567	ACTCGCTCTTCTCTCTCATTCTCCATCGATGGTGCCTATTACTTCGCGCAACTTACC	1626
Qy	561	TyrLeuArgGluArgIleThrArgTyrIleAspGluGluGluArgIleLysAlaThr	580
Db	1627	TACTTCGGAGACAGATTGGAAGTGGATCGATGAGAGAGAAATAAATGAGAACTAGC	1686
Qy	581	GlyGluThrSerLeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThr	600
Db	1687	GGAGATATCAGTCTACAATATATTGACTCGACATGAGTCTGTGGGAAACATCGACAG	1746
Qy	601	SerGlyLeSerMetLeuGluValLysLysIleThrCluArgArgGluLeuGlnLeu	620
Db	1747	AGCGGTATAGCATGTCTGAGAACTCAACAAATCTCTGGACGTAGAGAGCTTAAGCTA	1806
Qy	621	ValLeuValAsnProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsn	640
Db	1807	GTGATAGCAGACCGGAGCAGAGGTGATGAGAGAGTAAGCAAGTCCACCTTCATTGAG	1866
Qy	641	HisLeuGlyLysLysTyrIleTyrLeuThrValGluAlaValGlyAlaCysAsnPhe	660
Db	1867	AGCATTTGTTAAAGACCGATTATCTCACGGTGGCAGAACCGCTCGCAGCTTCGATTTC	1926
Qy	661	AsnLeuArgAlaSerLysThrAsnProLysLysAspGluThrGluGlyTyrAsnVal	680
Db	1927	ATGCTCCACACGGCTAAACCCGATCTCGCGGTACCGGAA-----TTTAAACAGTT	1977
RESULT 10			
AJ601439			
LOCUS			

DEFINITION	Brassica oleracea mRNA for sulfate transporter (LST3.2 gene).		
ACCESSION	AJ601439		
VERSION	AJ601439.1 GI:37998857		
KEYWORDS	LST3.2 gene; sulfate transporter.		
SOURCE	Brassica oleracea var. acephala (kale)		
ORGANISM	Brassica oleracea var. acephala		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.		
REFERENCE	1		
AUTHORS	Buchner,P., Stuijver,E.E., Hawkesford,M.J. and de Kok,L.J.		
TITLE	Analysis of the effect of sulphur nutrition and atmospheric H2S exposure on the expression of sulphate transporter genes in Curley kale		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2219)		
AUTHORS	Buchner,P.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-Oct-2003) Buchner P., Agricultural and Environmental Division, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UNITED KINGDOM		
FEATURES	Location/Qualifiers		
source	1..2219		
	/organism="Brassica oleracea var. acephala"		
	/mol_type="mRNA"		
	/variety="acephala"		
gene	1..2219		
	/db_xref="taxon:3713"		
5' UTR	1..71		
	/gene="LST3.2"		
	/gene="LST3.2"		
CDS	72..2024		
	/evidence=experimental		
	/gene="LST3.2"		
	/function="involved in plasma membrane sulfate transport"		
	/codon_start=1		
	/product="sulfate transporter"		
	/protein_id="CAE53112.1"		
	/db_xref="GI:37998858"		
	/translation="MCKRALQYQVETPPQPPQLSKNTLNEILFADDPFGKIRNES KTKKIELGRHVFPIEWARGYNLDYKSDVISGITIASLAIPQGISYAKLANLPPI LGLYSFPPMYVAIMGSKDLAVGTAVASLLTAAMLGKVSAYENPKYLHLAFTS TPFACLMQTCGLGLGLVELILSHAAIIGFNAGATVVCLOQLKGLGLSHPTHST VVSKRSIFSQSPVWRWESGLGCGCLFLPITIKYISKPKLFWISANAPLVSVIFG SUFYFLHAQHGIIQIIELEKGINPPSIHLVFTSPYVNLAKTGIITGVLAASGI AUYGRFAMYNKINIDGNKEMIAFGMWNIFGSSCYLTGPFSSRAVYNASCKTAVS NVVMAVAVATLLFTPLFFYPPLVLSIIIAAMLGLVDYEAMNMLRDKDFDFVC LSAPGVVGTTEIGLILSVGLSVLRLLFVGRPKIYVMGNIONTEPYRNIQYPOAT TLSSIIILHIDGPIYFANSSYLDRIGRWIDEEELKRSSENSLOYIILDLISAVNI DTSGIRMLVEEVNKILGRRLKLVIANPGALMKLKSFKPETIGKDWIHLTVAEAVS ACDFVLHTAKPDSPEKISREPEFNNV"		
5' UTR	2025..2219		
	/gene="LST3.2"		
	/evidence=experimental		
ORIGIN			
Alignment Scores:			
Pred. No.:	4,06e-172	Length:	2219
Score:	2274.50	Matches:	428
Percent Similarity:	81.50%	Conservative:	105
Best Local Similarity:	65.44%	Mismatches:	108
Query Match:	65.45%	Indels:	13
DB:	8	Gaps:	4
US-10-762-049-18 (1-680) x AJ601439 (1-2219)			
Qy	30	GluTyrProLeuGlyMetAsnAsnPheGluArgValHisGlnValGluValProProPro	49
Db	90	CAGTACCC-----CAAGTAGAAACCCCTCCACCA	119
Qy	50	GlnProPhePheLysSerLeuLysTyrSerLeuLysGluThrPhePheProAspPro	69
Db	120	CAACCTTTCTTGAAATCACTAAAGAACACTCTCAATGAGATCCTTTCCGCCGACGATCG	179

```
QY 70 LeuArgGlnPheLysAsnLys----ProAlaSerLysLysPheMetLeuGlyLeuGlnPhe 88
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
180 TTCGAAATAATCAGGAACGAATCGAAAAACGTCGAAAAAGATCGAGTTAGGTTGACACAC 239

QY 89 PhePheProIlePheGluTrpAlaProLysTyrThrPheGlnPheLysAlaAspLeu 108
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
240 GTGTTCCCGATCTTGGAAATGGGCTCGTGGTTACAATTTGACATCTACTTGAATCAGACGTT 299

QY 109 IleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLysLeu 128
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
300 ATCTCAGGATACGATAGCAAGTCTTGCTATTCTCAAGGAGTATAGTATGCTCAACTC 359

QY 129 AlaAsnLeuProProIleLeuGlyLeuTyrSerSerPheIleProProLeuIleTyrAla 148
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
360 GCTAAATCTCCCTCAATCTCTTGCTTTACTCGAGCTTCGTGCGCGGATGGTATACGCG 419

QY 149 MetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeuMet 168
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
420 ATAAATGGGAAGTTCAAAAGATTTAGCGGTGGCGACGCTAGCGGTGCGTCTGTTAAACG 479

QY 169 GlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHisLeu 188
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
480 CGCGCAATGCTTGGAAAAGAGGTAAAGCGCGGTGGAGAATCCAAAACTTTACCTTCACTTA 539

QY 189 AlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArgLeu 208
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
540 GCTTTTCACTCCATCTTCTTCGCGGACTCATCAACCTGCGCTTGGCTCTTTACGGTTA 599

QY 209 GlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAla 228
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
600 GGAATCTGTGTGAGATATTGTGCGCACGCGGCGATCATAGGGTTCAATGCTGGAGCAGCG 659

QY 229 ThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGly 248
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
660 ACGTGCTTTGCTCAACAACATGAAGGATTACTTGGTCTCAGTCACATTCACCCATTC 719

QY 249 AlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrpGlu 268
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
720 ACTGATGTCGTCTCGCTCTCTCTCCATCTCTTCTCAGTCTCCTCTGTGGAGATGGAG 779

QY 269 SerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSerLys 288
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
780 AGTGGGTGTACTAGGGTGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 839

QY 289 LysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeuGly 308
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
840 AAGAGACCAAGCTATTCTGGATATCTGCAATGGCTCCACTTGTCTGTCAATTTTGA 899

QY 309 SerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGluLeu 328
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
900 AGCTCTCTCGTTTACTTCTCCATGCCAGTCCATGGCATCCAAATATTGGAGATTTA 959

QY 329 LysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMetThr 348
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
960 GAGAAAGGAGTAAATCAACGTCATCAACACATCTGTTTTCACGTCACCTTATGTCATG 1019

QY 349 ThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAlaVal 368
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1020 CTGGCTCTTAAACCCGGCATTAATCACTGGAGTCTAGCTTGTGCTGAAGGAATTCAGTG 1079

QY 369 GlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAla 388
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1080 GGGAGAGCTTTGCATGTATAGATTAACATAGATGTTAAACAAAGATGATGCA 1139

QY 389 IleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyProPhe 408
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1140 TTTGGGATGATGAACAATTTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1199

QY 409 SerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleMet 428
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1200 TCACGTTCCGCGTGAACACTACACCGCATGTTGCAAAACCGCGGTATCAAAACGTTGGTATG 1259
```

```
QY 429 SerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrProLeu 448
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1260 CGGGTTGCAGTAGCAACGTTACTGTCTTAAACGCCATTATCTTTTATACGCCGTTA 1319

QY 449 ValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAlaAla 468
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1320 GTAGTTCTGTCTCGATCATCATCGCAGCATGCTTGGACTCGTCACTATGAAGCAGCG 1379

QY 469 IleHisLeuPheLysValAspLysPheAspPheValValCysMetSerAlaTyrIleGly 488
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1380 ATGAATCTTTGAGACTCGACAAGTTCGATTTCTTCGTATGTCCTCTCTCTCTCTCTCTCG 1439

QY 489 ValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeuArg 508
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1440 GTCGTTTTTGGTACCATTCGAGATTGGTCTCTCATCTCTCTGTGGGAATATCAGTCTCTG 1499

QY 509 ValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSerVal 528
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1500 TTACTGTGTTCTCGTGGGAAGACCAAGATTTATGTAATGGGAACATTCAGAACACAGAG 1559

QY 529 IleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeu 548
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1560 CCCTATAGGAACATTTGAACAGTACCCTCAAGCCACTACTCTCTCTAGTATCATCATCTCTC 1619

QY 549 GluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThrArg 568
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1620 CATATCGATGGTCTTACTTTACTTCGCCAACTCTAGCTACCTCGGAGATAGATTTGGAAG 1679

QY 569 TrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrVal 588
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1680 TGGATCGATGAGGAGGAGAGAAACGTAGAAAGAGTGAAGAAACCATCGGCAAGACATGGA 1739

QY 589 IleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGlu 608
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1740 ATACTTGATTTGAGTGTGTGGGAACATCGACAGTGGTATGAAGATGCTCCAGGAA 1799

QY 609 ValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProValSerGlu 628
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1800 GTCAACAAATACTCGGACGAAGAGACCTCAAGCTAGTAGTAGTAAACCCAGGAGCAGAG 1859

QY 629 ValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLysLysTrpIleTyr 648
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1860 TTGATGAAGAAGCTAAGCAAGTCCAAGTTCCAGAAACCATCGGCAAGACATGGAATTCAT 1919

QY 649 LeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerLys---Thr 667
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1920 CTCAGGTAGCAGAAAGCCGTCTCAGCCTGTGATTTCTGCTCTCCACACAGCTAAACCGGAC 1979

QY 668 AsnProLysLysAspGluThrGlu---GlyTrpAsnAsnVal 680
Db   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::   ::
1980 TCACCGGAGAAATCTCAGTGAACCGGAATTTAACACGTT 2021

RESULT 11
AR477457
LOCUS
DEFINITION
AR477457
ACCESSION
AR477457.1
GI:47234915
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
Unclassified.
1 (bases 1 to 1981)
AUTHORS
Allen,S.M., Falco,S.C. and Thorpe,C.J.
TITLE
Genes encoding sulfate assimilation proteins
JOURNAL
Patent: US 6696292-A 3 24-FEB-2004;
FEATURES
Location/Qualifiers
1..1981
source
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
```

Pred. No.:	2.39e-163	Length:	1981
Score:	2164.00	Matches:	403
Percent Similarity:	86.09%	Conservative:	86
Best Local Similarity:	70.95%	Mismatches:	79
Query Match:	62.27%	Indels:	0
DB:	6	Gaps:	0
US-10-762-049-18 (1-680) x AR477457 (1-1981)			
Qy	105	LysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSer	124
Db	8	GAGTCCGACCTGATCGCGGCATCACATCGCAGCCTCGCCATCCCGAGGGCATCAGC	67
Qy	125	TyrAlaLysLeuAlaAsnLeuProProlleLeuGlyLeuTyrSerSerPheIleProPro	144
Db	68	TACGCCAAGCTCGCCAACTGCGCGCGGTCTCGGACTCTACTCGAGCTTCGTGCGCGCG	127
Qy	145	LeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGly	164
Db	128	CTGGTGACCGCTGATGGGAGCTCAAGGACCTGGCGGTGGGACGGTGGCGTGGCG	187
Qy	165	SerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeu	184
Db	188	TCGCTGCTCATCAGCTCCATGCTCGCAGCGAGGTGTCGCCAGCGAGAACCCCGTGCTC	247
Qy	185	TyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGly	204
Db	248	TACTCGACCTCGCTTTCACCGCACCTTCTTCGCGCGGTCTTCACGCGCTCGCTCGGC	307
Qy	205	LeuPheArgLeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMet	224
Db	308	CTCCTCAGGTGGGCTTCATCGTGAGCTCTGTCGACGCGACGATCGTGCGGGTTCATG	367
Qy	225	GlyClyAlaAlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHis	244
Db	368	GCCGGCGCGGACGGTGGTGGCTGACAGAGCTGAAGGGCATGCTGGGCGCTCGTCCAC	427
Qy	245	PheThrHisGlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGlu	264
Db	428	TTCAACCACTCCACCGACGTGCTCGCTCATGGAATCGTCTTCAGCCACACACACAG	487
Qy	265	TrpArgTrpGluSerAlaValLeuGlyCysValPheIlePheLeuLeuSerThrArg	284
Db	488	TGGCGGTGGGAGAGCTGCTGCTCGCTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCT	547
Qy	285	TyrPheSerLysArgProArgPheThrValSerAlaMetAlaProLeuThrSer	304
Db	548	TTTCATCAGCAAGGCGTCCCAAGCTGTTCTGGATCTCGCGCGCGCGCGCTTACAGTCC	607
Qy	305	ValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGlyLysHisGlyValGluVal	324
Db	608	GTGCTGCTCGGAGCGTCTGTTGTAGCTCAGCAGCGCTGAAACCAACCGCATCGAGTG	667
Qy	325	IleGlyGluLeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSer	344
Db	668	ATCGGTACTCTGAAGAAAGCGCTGAATCCACCGTGGTGACAGCCCTGCAATTCTCACCG	727
Qy	345	ProTyrMetThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGlu	364
Db	728	CCCTACATGATGCTCGCGCTCAAGACTGGGATCATCACCGCGCTCATGCGCTCGCGAA	787
Qy	365	GlyIleAlaValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLys	384
Db	788	GGAATCGCGTGGGAGGAGCTTCGCCATGTTCAAGAACTTACCACATGACGCGCAACAAG	847
Qy	385	GluMetIleAlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThr	404
Db	848	GAGATGATCGCATCGGACGATGAACGCTCTCGGGCTGCTCACGCTCGTCTACTGACC	907
Qy	405	ThrGlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSer	424
Db	908	ACGGGGCCCTTCTCGCGCTCGCGGTGAATCAACCGCGGTGCAGGACGGCCATGTGCG	967
Qy	425	AsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHis	444
Db	968	AACGTGCTCATGCTCGCGGTGATGGTCACGCTGTTCTTCTCAGCGCGCTGTTCAC	1027
Qy	445	TyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAsp	464
Db	1028	TACAGCCCGCTGGTGGTGTCTGTCGCGCATCATGCTCTCCGCGATGCTGGGCTTGGTCGAC	1087
Qy	465	TyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheValValCysMetSer	484
Db	1088	TTCCGGGCGCGCTGTCACCTGTGGCGCTCGACAAGTGCATCTTCTGCGTCTGCGCGCGC	1147
Qy	485	AlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIle	504
Db	1148	GCGTACTCGGCGTCTGTTTCGCGACGCTCGAGGTGCGCTGCTGCTGCGCTCGCGCTC	1207
Qy	505	SerValLeuArgValLeuLeuPheIleAlaAArgProArgThrPheValLeuGlyValAsnIle	524
Db	1208	TCCTGCTCGCGTCTCTGTTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	1267
Qy	525	ProAsnSerValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGly	544
Db	1268	CCGGCACCATGTTGTCGCGAGGATGGACAGTACGCGCGCGCGCGCGCGCGCGCGC	1327
Qy	545	MetLeuIleLeuGluIleAspAlaProlleTyrPheAlaAsnAlaSerTyrIleuArgGlu	564
Db	1328	GTGCTGCTGCGCTGCGCTGCGCGCGCTTACTTTCGCCAACCGCGCTACCTCGGAGAG	1387
Qy	565	ArgIleThrArgTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSer	584
Db	1388	AGGATCTCGGTGGATCGACGAGGAGGAGCGCACCAAGAGCGCGCGCGCGCGCGC	1447
Qy	585	LeuGlnTyrValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSer	604
Db	1448	GTGCGGTAGTGTCTCTCGACATGGGTGGTGCATCGGTAGCATCGACACGCGCGCGC	1507
Qy	605	MetLeuGluGluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsn	624
Db	1508	ATGCTGGACGAGCTCAACAAAGTCTTGGACAGGAGGAGGAAATGCAGATCGTGTGGCGAAC	1567
Qy	625	ProValSerGluValMetLysLysLeuAsnLysSerLysPheGlnAsnHisLeuGlyLys	644
Db	1568	CGGGCAGCGAGATCATGAAAGAGCTGGACAGCTTCAAGGTGCTGGAGCAGATCGGCCAC	1627
Qy	645	LysTrpIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAla	664
Db	1628	GAGTGGGTGTTCCCGACGGTGGCGAGCGGTGGCGTCTGCTCGCACTACGTGCTGCATCG	1687
Qy	665	SerLysThrAsnProLysLysAsp	672
Db	1688	CACAAGCCGGGAATGGCCAAAGGAC	1711
RESULT 12			
ATP4F15/c			
LOCUS	ATP4F15	123209 bp	DNA linear
DEFINITION	Arabidopsis thaliana DNA chromosome 3, BAC clone F4F15.		
ACCESSION	AL049711		
VERSION	AL049711.3		
KEYWORDS	GI:6807566		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1		
AUTHORS	Alcaraz, J.P., Clabault, G., Cottet, A., Mache, R., Meves, H.W., Lemcke, K., Mayer, K.F.X., Quetier, F. and Salanoubat, M.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 123209)		
AUTHORS	EU Arabidopsis sequencing, project.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-JAN-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferstritz 18a, D-82152 Martinsried, FRG, E-mail:		

lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project  
Coordinator: Marcel Salanoubat and Francis Quetier, Groupement  
d'Interet Public, Centre National de Sequencage - GENOSCOPE; 2 rue  
Gaston Cremieux, BP191, 91006 Evry Cedex, France;  
http://www.genoscope.cns.fr  
On Jan 28, 2000 this sequence version replaced gi:4691223.  
Information on performance of analysis and a more detailed  
annotation of this entry and other sequences of chromosomes 3, 4  
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.

FEATURES  
Location/Qualifiers  
1..123209  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="3"  
536..4710  
/gene="F4F15.10"  
complement(join(536..721,804..880,977..1062,1145..1335,  
1471..1757,1849..1975,2172..2288,2597..2710,2950..3015,  
3169..3340,4033..4240,4365..4710))  
/gene="F4F15.10"  
complement(join(536..721,804..880,977..1062,1145..1335,  
1471..1757,1849..1975,2172..2288,2597..2710,2950..3015,  
3169..3340,4033..4240,4365..4710))  
/gene="F4F15.10"  
/notes="Contains prokaryotic membrane lipoprotein lipid  
attachment site AA369-379  
contains EST gb:T21459, Aa720146"  
/codon\_start=1  
/product="sulfate transporter (AtST1)"  
/protein\_id="CA41310.1"  
/db\_xref="GI:4678919"  
/db\_xref="GOA:Q9SV13"  
/db\_xref="UniProt/TREMBL:Q9SV13"  
/translation="MGEDYTFQGAELHRRHVTVEAPQPQFLKSLQYSVKETLFP  
DPPFRQKNASRKFVGLKFLPIFEWAPRNKLFKFKSDLIAGITIASLAPQGIS  
YAKLANLPILGLYSFVPLVYAVIGSSRDLAGVTAVASLLTGAMLSKEVDAEKDP  
KLYLHLPATATFAGVLEASLGI FRLGFIIVDFLSHATIVGFMGAATVVSLOQLKGI  
GLKHFTDSDTVLSVMSVFSOTHEWRESGLVGCGLPELLSTRVFSIKKPKFVVA  
MAPLSVILGSLVYFTHAERGVQVIGLKKGLNPLSGSDLIITSPYMSAVTKTGLI  
TGIIALAEVAVGRSFAMFNKNIDGNKEMIAFGMNNIVGSFSCYLITGPPRSVAVN  
YNAGCTAMSNVMAIVAMFTLLFLTPHYPVLVLSAIIISAMGLIDYQAAHLW  
KYDKFDLVMSAYGVWFGVSVEIGLVAVASIAKLIVSRPKTAVKGNTPNSMIY  
RTEQVPSRTPGLILIDAPIFVANASYLRIRIWIDEEERVKSGSSLOYI  
ILDMASVGNIDTSGISMMVEIKKVIDRRALKVLNPKGEVVKLTRSKFIGDLGKE  
WNFLTVEACSYMLHTFKTEPAKSNPEWNV"  
complement(536..721)  
/gene="F4F15.10"  
/number=1  
complement(722..803)  
/gene="F4F15.10"  
/number=1  
complement(804..880)  
/gene="F4F15.10"  
/number=2  
complement(881..976)  
/gene="F4F15.10"  
/number=2  
complement(977..1062)  
/gene="F4F15.10"  
/number=3  
complement(1063..1144)  
/gene="F4F15.10"  
/number=3  
complement(1145..1335)  
/gene="F4F15.10"  
/number=4  
complement(1336..1470)  
/gene="F4F15.10"  
/number=4  
complement(1471..1757)  
/gene="F4F15.10"

intron  
complement(1758..1848)  
/gene="F4F15.10"  
/number=5  
complement(1849..1975)  
/gene="F4F15.10"  
/number=6  
complement(1976..2171)  
/gene="F4F15.10"  
/number=6  
complement(2172..2288)  
/gene="F4F15.10"  
/number=7  
complement(2289..2596)  
/gene="F4F15.10"  
/number=7  
complement(2597..2710)  
/gene="F4F15.10"  
/number=8  
complement(2711..2949)  
/gene="F4F15.10"  
/number=8  
complement(2950..3015)  
/gene="F4F15.10"  
/number=9  
complement(3016..3168)  
/gene="F4F15.10"  
/number=9  
complement(3169..3340)  
/gene="F4F15.10"  
/number=10  
complement(3341..4032)  
/gene="F4F15.10"  
/number=10  
complement(4033..4240)  
/gene="F4F15.10"  
/number=11  
complement(4241..4364)  
/gene="F4F15.10"  
/number=11  
complement(4365..4710)  
/gene="F4F15.10"  
/number=12  
14327..15652  
/gene="F4F15.20"  
join(14327..14590,15098..15652)  
/gene="F4F15.20"  
/note="similarity to heat shock transcription factor HSF30  
- Peruvian tomato, PIR2:S25480  
Contains HSF-type DNA-binding domain signature AA69-93  
contains EST gb:R65204, Aa042693"  
/codon\_start=1  
/product="putative heat shock transcription factor"  
/protein\_id="CA41311.1"  
/db\_xref="GI:4678920"  
/db\_xref="GOA:Q9SV12"  
/db\_xref="UniProt/TREMBL:Q9SV12"  
/translation="MMNPFLPEGCDPPPPPOPEGLHENAPPLTKTFEMVDDPNTD  
HIYSWNRGTSFVVDLHSLFSTILLPRHFKSNFSSFIQLNTYGFRIEAEWSEFAN  
BEFLGQQLLNKIKRNPFTSPSSPHDACNLREKQVLMMEIVSRQOQOTKSY  
IKAMEQRLEGTERKORQWMSFLRAMQSPFLHOLLKQDKKIKELSDNESAKRKG  
SSMSLEVLIALEMQHGKQRNMLEEDHQLVVERLDDGFWEELLSDESLASTS"  
14327..14590  
/gene="F4F15.20"  
/number=1  
14591..15097  
/gene="F4F15.20"  
/number=1  
15098..15652  
/gene="F4F15.20"  
/number=2  
17211..18347  
/gene="F4F15.30"

```
gene      complement(join(17211..17351,17850..17927,18018..18178,
18272..18347))
/genes="F4P15.30"
CDS       complement(join(17211..17351,17850..17927,18018..18178,
18272..18347))
/genes="F4P15.30"
/notes="strong similarity to calmodulin - Tetrahymena
pyriformis (SGCS), PIRI:MCTE
contains EF-hand calcium-binding domain AA94-106;EF-hand
calcium-binding domain AA130-142
contains EST gb:R65197, R87041, Aa597913"
/codon_start=1
/product="putative calmodulin"
/protein_id="CA841312.1"
/db_xref="GI:4678921"
/db_xref="GOA:Q9S744"
/db_xref="UniProt/TrEMBL:Q9S744"
translators="MADAFTDEOIQBFYAFCLIDKSDGFIKELTKVMKSMGKNP
KASLOQMMSDVIDFGNGGITPDPLYIMAQNTSQSASDELLEVERFDRDGDGLIS
QLELGGMKDMGKIITAEBAHVMREADLDGDFLSFHFESKMIAASY"
exon      complement(17211..17351)
/genes="F4P15.30"
/number=1
intron    complement(17352..17849)
/genes="F4P15.30"
/number=1
exon      complement(17850..17927)
/genes="F4P15.30"
/number=2
intron    complement(17928..18017)
/genes="F4P15.30"
/number=2
exon      complement(18018..18178)
/genes="F4P15.30"
/number=3

Alignment Scores:
Pred. No.:      4,77e-150      Length:      123209
Score:          2026.50      Matches:      499
Percent Similarity: 41.40%      Conservative: 79
Best Local Similarity: 35.74%      Mismatches:   78
Query Match:    58.32%      Indels:       743
DB:             8           Gaps:         14

US-10-762-049-18 (1-680) x ATF4F15 (1-123209)
QY      24  MetGlySerValAspTyrGluTyrProLeuGlyMetAsnPheGluArgValHisGln 43
Db      4710  ATGGGACGGAGGACTACACATTCCTCAAGGAGCGGAGGAGCTTCACCGCCGCCAC 4651
QY      44  ---ValGluValProProGlnProPhePheLySerLeuLySerLeuLySerLeu 62
Db      4650  ACGGTGGAGGCTCCACAACCTCAACCGTCTTGAAGTCACTTCAGTACTCAGTGAAGGAA 4591
QY      63  ThrPhePheProAspAppProLeuArgGlnPheLyAsnLySerProAlaSerLySyphe 82
Db      4590  ACTCTGTGTTCCAGACGACCTTTTAGACAAATTTAGAACCAAAATGCATCAAGAAATTT 4531
QY      83  MetLeuGlyLeuGlnPhePheProIlePheGluTyrAlaProLySerTyrThrPheGln 102
Db      4530  GTGTAGGCTCAAAATCTCTCCCGAATTTGCAATGGGCACACGCTACAACTCTCAAG 4471
QY      103  PheLeuLySAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122
Db      4470  TTCTTCAAAATCAGATCTCATCGCGGAATCACCATCGTAGTCTGCCATCCCTCAGGGC 4411
QY      123  IleSerTyrAlaLySLeuAlaAsnLeuProIleLeuGlyLeu-Tyr----- 138
Db      4410  ATCAGTTACGCCAAACTTGCTAACTTGCCCCCAATTCCTGGCCCTTTGTAAAGCGTTATACA 4351
QY      138  ----- 138
Db      4350  TCTTTATTTTATATATTTTCCTTAAACAATTAAGGGTTTTTTTGTGTGTATATACACT 4291
```

```
QY      139  -----SerSerPh 141
Db      4290  ATTTTGAAGTAGCGTTTGTACTTATTTGTAACGTGTTTTTTTCGTTTAGATTGAGTTT 4231
QY      141  eIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVa 161
Db      4230  TGTACCGCATTTGGTATACGCGTCTAGGAGTTCAAGGACTTACGCGTGGAAACGGT 4171
QY      161  lAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValaAspProAsnGluAs 181
Db      4170  TCGGTTGCGTCTCTGTCACAGGTGCGATGCTGAGCAAAAGAAAGTTGATGCTGAGAAAGA 4111
QY      181  pProLySLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAl 201
Db      4110  TCCCTAAGCTTTACCTTCACCTTTCACCGCCACTTTTTTTCGCGCGCTTCGGAAGC 4051
QY      201  aAlaLeuGlyLeuPhe----- 206
Db      4050  CTCCTTGGAAAT-TTTCAGGTTCTTTTACTTTTATTTCTCCTTAATATTTTGTGCTT 3992
QY      206  ----- 206
Db      3991  TTAATATGTTTTTACTTTATGATAAAAGTGTGAGTAATATATTCGGTCCAATGCATGTGG 3932
QY      206  ----- 206
Db      3931  AAAAGATTCGCGCATAAAAAATATCAACCAATAGTAACTTTTTTCCTTACCTTTGATTCT 3872
QY      206  ----- 206
Db      3871  AAGATCTAAAAATTCATAAGCACTATGTACTATGTACTATATAGAGAGTGAATTCGTCAT 3812
QY      206  ----- 206
Db      3811  CGCGAGTTACGTCACGTGAAACAAGATACAAGAGCTGAGACAAACAACAATAGGACAA 3752
QY      206  ----- 206
Db      3751  AGCTAGATAACTATTAGAAAAATCACACCGTGGTATGCTCGGTTTAATAATACATATACTAC 3692
QY      206  ----- 206
Db      3691  ATGTAACAGTGGAGAAATCATACTAGAAATATGTAATATCTTTACTGTAAGTTGGGTATA 3632
QY      206  ----- 206
Db      3631  ATTAAATACATATCTGCCCAAGTTTATTTTGATTCGCCGTATCATATAATCATAGTGTGATT 3572
QY      206  ----- 206
Db      3571  GATTCAAAAGCACAGAGAGATATATATATATTATAGGTAGAAATCTTTTATTTTGGTTC 3512
QY      206  ----- 206
Db      3511  CACAAAATTTCCGTATATCTAGCCAGAACTGGATTGAAGCATTAAGGTATTAACTAAGTT 3452
QY      206  ----- 206
Db      3451  GGTGAGTTTGATGATGATGTGAGTTTAAAGTGCACAACTGATAAATTTGATTTTAGTAG 3392
QY      207  -----ArgLeuGlyLe 210
Db      3391  CTTTAAAAATAAAAATTAAGAGAAGAAATGATGAAAAAGGAAATTCAGAGTTAGGGTT 3332
QY      210  urleValaAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyValaAlaThrVa 230
Db      3331  CATAGTGGACTTTCTATTCGCATGCAACGATAGTAGGATTTCATGGGAGGAGCAGCAGGT 3272
QY      230  lValCysLeuGlnGlnLeuLySerIleLeuGlyLeuGluHisPheThrHisGlyAlaAs 250
Db      3271  GGTGAGTCTGCAACAGCTTAAGGGTATTTTCGGACTTAAACATTTTACAGACTCTACCGA 3212
```





Db	1053	GTGGATTGATGAAGAGGAAGAGAGAGTAAACAATCAGGAGAGACGCTTACAAATATAT	994	/mol_type="genomic DNA"	
Qy	588	l1le1leAaspMet-----	592	/db_xref="taxon:3702"	
Db	993	TATACCTGATATGTCAGGTAATGCATATGATCTTGATTTTCTAACACTAGCTAATGTTG	934	complement(join(922..1362,1541..>1786))	mRNA
Qy	593	-----	596	/product="hypothetical protein"	
Db	933	TTTTTCGTTGAGCGCTTACTTAGTTACTCTGTCGCTGTTTGTTTTAAACAATCAGCTGTTGG	874	complement(join(922..1362,1541..1786))	CDS
Qy	596	yAn1leAaspThrSerGly1leSerMetLeuGluGluVal1leThrGluAArg	616	/note="hypothetical protein; ORF1; similar to sequence in GenBank Accession Number AC002330"	
Db	873	TAATATCGACACACGCGTATAGCATATGATGCGAAATTAAGAAGTCATTGACAGGAG	814	/codon_start=1	
Qy	616	gGluLeu-----	618	/product="unknown"	
Db	813	AGCGTTAAAGGTAACGTAGTCGCTGTAATAACATTTTGTAGATGATGAACGCGTTCTTT	754	/protein_id="AAC14403.1"	mRNA
Qy	619	-----	629	/db_xref="GI:3068703"	
Db	753	CTTGTTAATATCTTTCTTTTAAATCCACAGTGTGATGTTGCAATCCAAAAGGAGAGG	694	/translation="MILYLTVEEMDINRSLNPNFNAFFLSKTHVYQEKLPQGFPPA	
Qy	629	alMetLysLysLeuAenLysSerLysPhe---GlnAsnHisLeuGlyLysLysTrpIleT	648	SVFFCWETAPGPKNSHLKNDASETPRLKLPGRLEKYFQHTNQDAVDLSUTQAI	
Db	693	TCGTGAGAAATTAACACGATCCAAATTCATCGGTGATCATTTGGCGCAAGAGTGGATGT	634	DNVELPKDSVESDDRSVSGSDNGYLTMWSESTDESPSYIIERFLPDAALAAV	
Qy	648	yrLeuThrValGluGluAlaValGlyAlaCyAsnPheAsnLeuAArgAlaSerLysThrA	668	TSAAQRKKKLSYLSGETVRQSCFSPKACGLHVLLEPSTKHRIICGVKNAFSPSSHHV	
Db	633	TCCTAACGGTAGAGAGACAGTGGAGCGTTGTAGCTACATGCTTCACACGTTTAAACCG	574	LQPKFTTKDN"	
Qy	668	snProLysLysAepGluThrGluGlyTrpAsnAsnVal	680	complement(join(<2709..2935,3332..3641,3930..4221,4329..>56376))	
Db	573	AACCG---GCCTCCAAAACGAGCGCTTGGACAACGTA	539	/product="hypothetical protein"	
RESULT 13	AF049236/c			complement(join(2709..2935,3332..3641,3930..4221,4329..6376))	
LOCUS	DEFINITION			/note="hypothetical protein; ORF2; similar to sequence in GenBank Accession Number AC002330"	
DEFINITION				/codon_start=1	
KEYWORDS				/product="unknown"	
SOURCE				/protein_id="AAC14404.1"	
ORGANISM				/db_xref="GI:3068704"	
REFERENCE				/translation="MOHNLFTTMSLKLAECKTGVVALNPSATPPPPPGNSSTGSG	
AUTHORS				PVTDLLEQIDFLCFLVDLVEKMAQVRRICQSFERSGAYLEQCAIFRIGISDPLK	
TITLE				RRSLRSPQHAADVHAKVVLASWSLSRSGDELIGTTSMDCCGRNLEKPKATLVSGDPC	
JOURNAL				ESVYDPCVCSGASRSEMMNDECSQSDVMDYFCIDEBEVCVRKYIASLRPFKFA	
MEDLINE				DELKACDSHLAHLNLSIDEAMLLIEYGLSEAAVLLVAACLOVFLRELPSMMENPNVI	
PUBMED				KIFCSAEGRELASGHASFLLYFFLSQIAMEDDMKSNNTVTLMLERLVECAVSWEKQ	
REFERENCE				LAYHGVVLMERKEYKDAQRFNAAVEAGHLYSLGVARTKFKRDHRYSAIKIINSL	
AUTHORS				ISDHKATQMHQERSLYCSGKEKLLDLDATFEDPTLTTPFYKFAVALVEENQFGAAI	
TITLE				VELLRPLAQMSQADCMOLYDRWSVDIGSLAVVHHMLANDPKGLLAFROSLLLLH	
JOURNAL				RLNCQKAMRSILRLARNHKSSEHERLAVVEGWLVDYTGHRREALAKAESISIORSPA	
MEDLINE				FLFKAYALDSTLDPDSSNYVQLLQALAKCPSGDLKRGQALNNLGSVYVDCBKLDLDA	
PUBMED				ADCVYNALTIKHTRAHOGRLVHLLKNQKAAVDEMTKLIEKQANNASAYEKKSEYCD	
REFERENCE				REMAQSDICLALQDLPRTYPPYRAAVALMDHDKHSEAIIDELSRATSELPDQLLHLR	
AUTHORS				AAFYDSMGEGASAIKDCEAALCIDPGHADTLELHYHKAEPNDQK"	
TITLE				complement(join(9320..9605,9694..9840,9918..10174,10350..10662))	
JOURNAL				/product="hypothetical protein"	
MEDLINE				complement(join(9408..9605,9694..9840,9918..10174,10350..10557))	
PUBMED				/note="hypothetical protein; ORF3; similar to Arabidopsis thaliana clone 240M377 sequence found in dbEST Accession Number N65910"	
REFERENCE				/codon_start=1	
AUTHORS				/product="unknown"	
TITLE				/db_xref="GI:3068705"	
JOURNAL				/translation="WMNSTSESEVVRPGMLVQRDDTASSDHPKLPQDPDSASAAF	
MEDLINE				QAQTVITVSHSGSSHHDLHLSHATFGDKVKALVQKTGCTEASELKLFRGVRDDAEOL	
PUBMED				QAAQVYDASKVLVVEDTNKRVQPPVVTTEMKETAANAVNTGEVVDKLSDRVAALGE	
REFERENCE				VAVNGQTVAVREFDMAELRLQLKLDGIEAGDQKAVQKAEVRIQNLQAEVADLQ	
AUTHORS				KARCSNPFVDQSKAAAVSTEWESFGVSGSLNPPPPASPSANVTQDWEKFD"	
TITLE				complement(12025..13047)	gene
JOURNAL				/gene="AtG1"	
MEDLINE				complement(<12272..>13042)	mRNA
PUBMED				/gene="AtG1"	
REFERENCE				/product="putative transmembrane protein G1p"	
AUTHORS				complement(12272..13042)	CDS
TITLE				/gene="AtG1"	
JOURNAL				/note="similar to Bradyrhizobium japonicum CycJ; SwissProt Accession Number P45401"	
MEDLINE				/codon_start=1	
PUBMED				/product="putative transmembrane protein G1p"	
REFERENCE				/protein_id="AAC14406.1"	
AUTHORS				/organism="Arabidopsis thaliana"	
TITLE				1..60000	FEATURES
JOURNAL				Location/Qualifiers	source
MEDLINE					
PUBMED					



gene	/db_xref="GI:3068706"		mRNA	join(22568..22766,22902..22976,23067..23194,23297..23388,23469..23526,23639..23695,23775..23808,23947..24027,24107..24213,24563..24649,24731..24808,24982..25046,25268..25366,25516..25615,25979..26087,26196..26268,26601..26694,26792..26875,27183..27249,27344..27625)/gene="AtG5"	
	/product="putative transmembrane protein G5p"				
gene	complement(13443..16043)		mRNA	/gene="AtG2"	
	complement(join(13443..13716,13805..13978,14067..14158,14245..14381,14467..14682,15169..15322,15415..15564,15660..15744,15854..16043))				
CDS	/product="putative nuclear DNA-binding protein G2p"		CDS	/product="putative transmembrane protein G5p"	
	complement(join(13640..13716,13805..13978,14067..14158,14245..14381,14467..14682,15169..15322,15415..15564,15660..15744,15854..15947))				
gene	/gene="AtG2"		mRNA	/note="similar to yeast Sac1 protein"	
	/note="similar to Mus musculus p38-2G4 protein: SwissProt Accession Number P50580"				
gene	/codon_start=1		mRNA	/codon_start=1	
	/product="putative nuclear DNA-binding protein G2p"				
gene	/db_xref="GI:3068707"		mRNA	/protein_id="AAC14407.1"	
	/translation="MSSDDEKELSTSPVVTYKYSAAEIVNKALQVLAECKPKAKIVDICEGDSFIKEQTASMYKNKKIERGVAFPTCISVNNTVGHFSPLASDSVLKNTDVTAEAKQVAAADCKIVEGLSHQLKHQVLDGNKVLSVSPETTVDBVEFEENEYVIDIVASTGDKPKLDEKQTTIYKKDESNNYOLNKKASRFIISIKQNPMPFTARSLERKARGLVECVNHHLOPPVLYYKFGDFVAQIKFTYLLMNGSDRIISHTLQLEPKTITDEPFIKGLWLGILKKKGGKKKAQKAGKEASTEAEPEMDASSNAQE"				
gene	17135..18031		mRNA	US-10-762-049-18 (1-680) x AF049236 (1-60000)	
	/gene="ATEM1"				
CDS	join(17135..17314,17500..18031)		mRNA	QY 24 MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln 43	
	/gene="ATEM1"				
CDS	/product="Em1 protein"		mRNA	DB 58057 ATGGGACGCGAGCATACACATTCCCTCAAGGAGCGGAGGAGCTTCCACGCGCCACAC 57998	
	/gene="ATEM1"				
CDS	join(17197..17314,17500..17840)		mRNA	QY 44 ---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu 62	
	/codon_start=1				
gene	/product="Em1 protein"		mRNA	DB 57997 ACGGTGGAGGCTCCACACCTCAACCGTCTTGAAGTCACTTCAGTACTCAGTGAAGAA 57938	
	/protein_id="AAC14408.1"				
gene	/db_xref="GI:3068708"		mRNA	QY 63 ThrPhePheProAspAspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe 82	
	/translation="WASKQLSREELDEKAKQGETVVPQGTGGHSLAQPHLAEGRSKGQTRKEQLHGEGYQEIIGHKGGEARKEQLHGEGYQEMHGKEARKEQLHGEGYQEMGHKGBARKEQLHGEGYQEMGRKQGLSTMEKSGGERABEEGIEIDESKFTNK"				
gene	complement(18488..21370)		mRNA	DB 57937 ACTCTGTTTCCAGACGACCCCTTTAGACAAATTTAAGAACCAAAATGCATCAAGAAATTT 57878	
	/gene="AtG4"				
CDS	complement(join(18488..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21370))		mRNA	QY 83 MetLeuGlyLeuGlnPhePhePheProIlePheGluTyrAlaProLysTyrThrPheGln 102	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	DB 57877 GTGTTAGCGCTCAAAATACTTCTCCCGATTTTCGAATGGCCACCGCTACCAATCTCAAG 57818	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	QY 103 PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly 122	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	DB 57817 TTCCTCAATCAGATCTCATCGCCGGAATCACCATCGTAGTCTCGCCATCCCTCAGGCG 57758	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	QY 123 IleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeu-Tyr----- 138	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	DB 57697 TCTTTTATTTTATTATTTCCTTAACAATTAACGGGTTTTTTTGTGTGTATATACACT 57638	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	QY 139 -----SerSerPh 141	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	DB 57637 ATTTTGAAGTAGCGTTGTACTTTATTGTAAGTGTTCGCGTTTATAGATTCGAGTTT 57578	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	QY 141 eileProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVa 161	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	DB 57577 TGTACCGCATTTGGTATACGCGGTGCTGAGGAGTTCAAGGAGCTTAGCGGTGGAAACGGT 57518	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	QY 161 AlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAs 181	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	DB 57517 TCGCGTTCGGTCTCTGTTGACAGGTGCGATCGTAGCAAGAAAGTGTGCTGAGAAAGA 57458	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	QY 181 pProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAl 201	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	DB 57457 TCCTAAGCTTTACCTTCACCTTGCTTTCCGCCCACTTTTTCGCCGGCGTTCCTCGAAGC 57398	
	/gene="AtG4"				
CDS	complement(join(18680..18742,18838..18885,18970..19046,19410..19554,19644..19719,19798..19859,19956..20039,20114..20179,20253..20339,20428..20513,20596..20683,20868..20953,21037..21081,21163..21313))		mRNA	QY 201 aAlaLeuGlyLeuPhe----- 206	
	/gene="AtG4"				

Db	57397	CTCTCTGGAAATTTTCAGGTTCTTTTTTACTTTTATTTCTCCTTAATTAATTTTGTGCTTT	57338
Qy	206	-----	206
Db	57337	TAATATGTTTACTTATGATAAAAGTGTGAGTAATATAATTCGGTCCAATGCATGTGGGA	57278
Qy	206	-----	206
Db	57277	AAAGATTCGGCATAAAAAATATCAACAAATAGTAACCTTTTCCCTTACCTTGATCTTGA	57218
Qy	206	-----	206
Db	57217	AGATCTAAAAATTCATAGCACTATGTACTATGTACTATATGAGAGTGAATCGGTCATC	57158
Qy	206	-----	206
Db	57157	GCGAGTTACGTCAGTGAACAAAGATACAAAGCTGAGACAACAAACCAATAGGACAAA	57098
Qy	206	-----	206
Db	57097	GCTAGATAACTATTAGAAAAATCACACCGTGTATGCTCGGTTTAAATACATATACTACA	57038
Qy	206	-----	206
Db	57037	TGTAACAGTGGAGAACTACTAGAAATGTAATTAATTAATTAATTAATTAATTAATTAAT	56978
Qy	206	-----	206
Db	56977	TTTAATACATATCTGCCAAGTTTATTTGATCCCCGTATCATATAATCATAGTAGTGATGA	56918
Qy	206	-----	206
Db	56917	TTCAAAAGCACAAGAGAGATATATAGTATATTATAGTGAGAACTTTTATTTTGGTTCCA	56858
Qy	206	-----	206
Db	56857	CAAAATTCGGTATATCTAGCCAGAACTGGATTGAAGCATAAAGGTATTAACTAAGTTGG	56798
Qy	206	-----	206
Db	56797	TGAGTTTGATGATGTGAGTTTAAATAGTGCAAACTGATAAATTGATTTGTTAGTACT	56738
Qy	207	-----	207
Db	56737	TTAAAAATAAAATTAAGAGAAGAAAGAAATGATGMAAAGGAAATTCAGGTTAGGGTTCA	56678
Qy	211	leValAspPheLeuSerHisAlaThrIleleGlyPheMetGlyGlyAlaAlaThrValV	56618
Db	56677	TAGTGGACTTTCTATCGCATGCCAACGATAGTAGGATTCATGGAGGAGCAGCAGCTGG	56618
Qy	231	alCysLeuGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAlaAspI	56618
Db	56617	TGAGCTGCAACAGCTTAAAGGTATATTCCGACTTAAACATTTTCACAGACTCTACCGATG	56558
Qy	251	leIleSerValMetArgSerValPheThrGlnThrHisGlu	56558
Db	56557	TTATCTGTGTCATGCGGTTCCGGTTTCTCCCAAACTCACGAGGTGCTCTCATTTCCGGTGA	56498
Qy	264	-----	264
Db	56497	TGTTGCTAATACATCCAAAATATCATTTTCAATGTGATGATGTTCTCTCAATTAATATAC	56438
Qy	264	-----	264
Db	56437	TATCAATATTCATAGTATTATTAACAGTAAACGCTCTAAACATGTGTTTGTGTTGTTG	56378
Qy	265	-----	265
Db	56377	TTTTTGTGTGTAGTGGAGATGGAAAGTGGCGTTCTTGGATGTGGTTTCTATTTCTTCTTC	56318
Qy	280	eulSerThrArgTyrPheSerIys	56258

---

Db	56317	TTCTCTCCACCAGATATTTTCGT-AAGATTATTAATTGTATACCTTTTATATATCAATTTAAA	56259
Qy	288	-----	288
Db	56258	TAAGTCACTAAATTAACACATTTTATTTTCGAATAGCAGCTACAATGATCTTTTCGAATAAT	56199
Qy	288	-----	288
Db	56198	TCCATTACTTTTACTCTTCGTCTTTTCAATCTTTTCTAGAAAGAATAAAAAAGTATAC	56139
Qy	288	-----	288
Db	56138	GATATGGTGGACTGAAAAATGTAATAACTACACAACTCAAAACGTTTATTAATTCGTGTTAA	56079
Qy	289	-----	289
Db	56078	TGAATGAAATGATGATGACAGAGCATCAAGAAACCAAAATCTTTTGGTGGCGCGATGG	56019
Qy	300	laProLeuThrSerValIleLeuGlySerLeuValTyrPheThrHisAlaGluLysH	56019
Db	56018	CTCCTTTGACCTCAGTGATCTTTGGAAAGTCTCTTGGTTTACTTCACTACGCTGAGAGAC	55959
Qy	320	isGlyValGluValIle	325
Db	55958	ATGGTGTTCAGTGGTATTACTCAATTTCTAGNAATCTCCTTATCTGATGATAAATAATCA	55899
Qy	325	-----	325
Db	55898	AATAATAAATTAATTTCAAAGTCAAGACCAAGAGAGATAACGTTGATGATTTTACCGAATCT	55839
Qy	325	-----	325
Db	55838	TAAGAACTTATATATTTTAAAGAAAGAAAGTTAGGTAGAATAATTTATGGTAAGA	55779
Qy	325	-----	325
Db	55778	AAAGAACAAAGACATGAAGTTGAAGTAGTAAAGAAATGTGGAAAAAGAGTGGGCTC	55719
Qy	325	-----	325
Db	55718	ATGGGCTTTTGTAGTTTGTTCGAAAGGAGAGGAATTAATCGAATAAAGAAATTAAGG	55659
Qy	326	-----	326
Db	55658	TATACCGATTGTGATGTGCAGATAGGGAC-CTGAAGAAAGGGTTGAATCCACTCTCCG	55600
Qy	337	eulThrAsnLeuValPheValSerProTyrMetThrThrAlaValLysThrGlyIleValV	55540
Db	55599	GTTCTGATCTCATCTTTTACTTCCCTTACATGTCCACAGCTGTCAAAACTGGGCTCATCA	55540
Qy	357	alGlyIleIleSerLeuAla	363
Db	55539	CTGGCATCATTTGCTCTCGC-TGTAAGCCCATTAATCCCTTTTCTTATTATTAGGCTTTTGTAT	55481
Qy	363	-----	363
Db	55480	AGTTATCTTATTTATTACCAAGTGTAAATTAACAAATTAATACTAAGCTTTTATACCTAAGGC	55421
Qy	363	-----	363
Db	55420	TTAATATCAACACACACATATGGGAATATGATTTTCAATGGTGAAGAAAGTGAACACAAA	55361
Qy	364	-----	364
Db	55360	ACTCATCAAAATCAAAATCAAAATTCATGTGTAAATATGTAGGAAGGAGTAGCAGTGGGAGGA	55301
Qy	371	erPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIleAlaIleGlyT	55241
Db	55300	GTTTTGCGATGTTTCAAGAACTCAACATAGACGGGAACAAAGAGATGATAGCGTTTGGAA	55241
Qy	391	hrMetAsnValValGlySerPheThrSerCysTyrLeuThr	404
Db	55240	TGATGAATCATCGTTGGTTCTCTTACATCTTTGTTTACCTCAACACCGGTACCTACACACTCA	55181

```
QY 404 ----- 404
Db 55180 CACATAGTCACATATGGTTATTAACTATGAAGAAAAAGCTTATATATTAAATTACGACG 55121
QY 405 ----- Thr-GlyProPheSerArgSerAlaValAsnTyrAenAlaGlyCysLys 420
Db 55120 TACCATACTATACAGGACCATTTTCAAGGTCGCGAGTGAACATACAAACGCGGGTTGCAAG 55061
QY 421 ThrAlaAsnSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLeuThr 440
Db 55060 ACCGCAATGTCACACATAGTAGGCGATGGCGGTATGTCACCTCTTCCTTCCTCAC 55001
QY 441 ProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMetLeu 460
Db 55000 CCGCTTTTTCACATACACACCATCGTCGTCCTCTCTGCCATCATCATCCGCAATGCTC 54941
QY 461 GlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLysValAspLysPheAspPheVal 480
Db 54940 GGACTCATTTGACTATCAAGCTGCATCCATCTCTGGAAAGTTGACAAAGTTCCGACTTCTC 54881
QY 481 ValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuValIle 500
Db 54880 GTCTGATGAGCCCTACGTTGGGTCGTATTCGGCAGTGAGATTGCACTCGTCGTA 54821
QY 501 Ala----- 501
Db 54820 GCGGTTAGTATATACTTTATATGCTGAATGACCTAGATGCCATGCTCTTTTTC 54761
QY 501 ----- 501
Db 54760 CTTCTATTTCGCTTGGCATTTTAAACATTCGCAATATTATTAATTTCTAAAATGTTTTT 54701
QY 502 ----- IleValIleSerValLeuArgValLeuLeuPheIleAlaArg 515
Db 54700 TTTTTCATGCTGATGAAGTGGCGATATCTATAGCAGGTTGTCTGTTGTGTCGAGG 54641
QY 516 ProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGluHis 535
Db 54640 CCAAAACTGCGGTGAAGGAAACATACCAACAGCATGATCTATAGGAACACTGAGCAG 54581
QY 536 TyrGlnAsnAlaLysHisValProGlyMetLeuIleGluIleAspAlaProIleTyr 555
Db 54580 TACCCATCATCAAGAACCGTTCTGTGATTTCTCATCTTGAGATTGATGCTCCCATCTAC 54521
QY 556 PheAlaAsnAlaSerTyrLeuArgGlu----- 564
Db 54520 TTTGCTTAATGCCAGTTACTTGGTGAAAGGTAACTAACCTCATTCGTATTATACATGCC 54461
QY 565 ----- ArgIleThrAr 568
Db 54460 AAAGATTGAAGATGCTTTTGAATAATGTTTACTTTTCAATAAACTACAGATCATAG 54401
QY 568 gTrpIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyrVa 588
Db 54400 GTGGATTGATGAGAGAGAGAGAGAGTTAAACAATCAGGAGAGCAGCTTACAATATAT 54341
QY 588 lIleIleAspMet----- 592
Db 54340 TATACTCGATATGTCAGGTAATGCATATGATCTTGATTTTCTTAACACTAGCTAATGTT 54281
QY 593 ----- SerAlaValG1 596
Db 54280 GTTTTCGTCGAGGCTTACTTAGTTACTCTGTGCGTGTGTTGTTAAACAATCAGCTGTCG 54221
QY 596 yAsnIleAspThrSerGlyIleSerMetLeuGluValLysLysIleThrGluArgAr 616
Db 54220 TAATATCGACACAAGCGGTATTAGCATGATGCTGGTGAATAATTAAGAAAGTCAATTGACAGGAG 54161
QY 616 gGluLeu----- 618
Db 54160 AGCGTTAAAGGTAACTAGTGGTGTAATAATACATTTTATAGGATATGTAACGCGTTTCTT 54101
```

```
QY 619 -----GlnLeuValLeuValAsnProValSerGluV 629
Db 54100 CTTGTTAAATATCTTTCTTTTAAATCCACAGTTGGTATTGTCAAATCCAAAGGAGAGG 54041
QY 629 alMetLysLysLeuAsnLysSerLysPhe---GlnAsnHisLeuGlyLysLysTyrIleT 648
Db 54040 TCGTGAAGAAATTAACCATCATCAATTCATCGGTGATCATTTGGGCAAGAGTGGATGT 53981
QY 648 yrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerLysThrA 668
Db 53980 TCTTAACGCTAGGAAGACGAGCTGGAGCTGTAGCTACATGCTTCACACGTTTAAACCG 53921
QY 668 snProLysLysAspGluThrGluGlyTyrAsnAsnVal 680
Db 53920 AACCG---GCCCTCAAACACGAGCCTTGGACAACAGTA 53886
RESULT 14
AB012048
LOCUS
DEFINITION
AB012048 7012 bp DNA linear PLN 14-FEB-2004
Arabisidopsis thaliana gene for sulfate transporter, complete cds,
clone:AST12.
ACCESSION
AB012048
VERSION
AB012048.1 GI:2967455
KEYWORDS
sulfate transporter.
SOURCE
Arabisidopsis thaliana (thale cress)
ORGANISM
Arabisidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1
Takahashi,H., Sasakura,N., Kimura,A., Watanabe,A. and Saito,K.
Identification of two leaf-specific sulfate transporters in
Arabisidopsis thaliana
Plant Physiol. 121, 686 (1999)
REFERENCE
2 (bases 1 to 7012)
Takahashi,H., Sasakura,N. and Saito,K.
Direct Submission
Submitted (06-MAR-1998) Hideki Takahashi, Chiba University, Faculty
of Pharmaceutical Sciences; 1-33, Yayoi-cho, Inage-ku, Chiba, Chiba
263-8522, Japan [E-mail:htaka@p.chiba-u.ac.jp, Tel.81-43-290-2906,
Fax:81-43-290-2905]
FEATURES
Location/Qualifiers
source
1..7012
/organism="Arabisidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="AST12"
/ecotype="Columbia"
join(1510..1855,1980..2187,2880..3051,3205..3270,
3510..3623,3973..4053,4250..4376,4468..4754,4890..5080,
5163..5248,5346..5422,5505..5690)
/codon_start=1
/product="sulfate transporter"
/protein_id="BAA25175.1"
/db_xref="GI:2967456"
/translation="MGTEDTYTPQGAELRRHHTVEAPQPPFLKSLQYSVKETLFP
DDPRFQKQNASRKFVLGKLYELPIFEWAPRYNLKFKFSKLSLGIATIASLAIPGDIS
YAKLANLPILGLKSSFPPLVYAVLGSRDLAVGTVAASLLTGAMLSKEVDASKDP
KLYLHLATATFFAGVLEASIGIFRLGFIVDPLSHATVGFMGGAATVVSUQLKGIF
GLKHFDSTDISVMSRVSQSTHERWESGLVGCGLFLLSTYFSTFKKPPFWAA
MAPLTSVLIGLLVYFTHAERHGVQVSDLIFTSPYMTSTAVKTLGITGIIALAEVAV
GRSPAFKNNYINDGNKEMIAFGMNIIVGSFTSCYLTGTFPFRSAVYVAGCKTASNI
VMAVMTFLPLPLPHVTLVLSAIIISAMGLIDYQAIHLWKVDKDFPLVCM5
AYGVVFSIGLIVAVASISARLLLFWSPKTAVKGNINSMIYRNTVEQYPSRTV
PGILLEIDAPIYFANASYLRIRIRIDEERVKQSGSLQYIILDM5AGNIDT
SGISNMVIRKVIDRRALKVLNSNPKGEVVKLTFRSKFIGDHLGKEMWFLTVGEAVEA
CSYMLHTEKTPSPAKNEPWNV"
ORIGIN
Alignment Scores: 1.57e-150 Length: 7012
Pred. No.: 2012.50 Matches: 500
Score:
```

Percent Similarity:	41.39%	Conservative:	79
Best Local Similarity:	35.74%	Mismatches:	77
Query Match:	57.91%	Indels:	745
DB:	8	Gaps:	14
US-10-762-049-18 (1-680) x AB012048 (1-7012)			
Qy	24	MetGlySerValAspTyrGluTyrProLeuGlyMetAsnAsnPheGluArgValHisGln	43
Db	1510	ATGGGCACGGAGGACTACACATTTCCCTCAAGAGCGGAGGAGCTTCACCGCCGCCACAC	1569
Qy	44	---ValGluValProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGlu	62
Db	1570	ACGTGGAGGCTCCCAACCTCAACCTTCTTGAAGTCACCTCAGTACTCAGTGAAGAA	1629
Qy	63	ThrPhePheProAspAppProLeuArgGlnPheLysAsnLysProAlaSerLysLysPhe	82
Db	1630	ACTCTGTTCCAGACGACCTTTTAGACAATTTAAGAACCAAAATGCATCAAGAAATTT	1689
Qy	83	MetLeuGlyLeuGlnPhePhePheProIlePheGluTyrAlaProLysTyrThrPheGln	102
Db	1690	GTGTAGGCTCAAACTTCTCCCGATTTTCCGAATGGGCACACCGCTACAAATCTCAAG	1749
Qy	103	PheLeuLysAlaAspLeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGly	122
Db	1750	TTCTTCAAAATCAGATCTCATGCGGGAATCACCATCGCTAGTCTCGCCATCCCTCAGGGC	1809
Qy	123	IleSerTyrAlaLysLeuAlaAsnLeuProProIleLeuGlyLeu-Tyr-----	138
Db	1810	ATCAGTTACGCCAAACTTGCTTACTTGCCCCCAATTCCTGGCCTTGTGAAGCGTTATACA	1869
Qy	138	-----	138
Db	1870	TCITTTATTTTATTTATTTTCCTTAACAATTAACGGGTTTTTTGTGTGTATATACACT	1929
Qy	139	-----SerSerPh	141
Db	1930	ATTTTGTGAAGTAGGTTTGATCTATTGTAACTGTTTTTTTTCGTTTAGATTTCGAGTTT	1989
Qy	141	eIleProProLeuIleTyrAlaMetMetGlySerSerArgAspLeuAlaValGlyThrVa	161
Db	1990	TGTACCGCCATGTTGATACGCGGTGCTAGGAGTTCAGGGACTTAGCGGTGGAAACGGT	2049
Qy	161	lAlaValGlySerLeuLeuMetGlySerMetLeuSerAsnAlaValAspProAsnGluAs	181
Db	2050	TGCGGTTGCGCTCTGTGTGACAGGTGCGATGCTGAGCAAAAGAGTTGATGCTGAGAAAGA	2109
Qy	181	pProLysLeuTyrLeuHisLeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAl	201
Db	2110	TCCTAAGCTTTACCTTCACCTTGCTTTCACCGCCACTTTTTTCGCGCGGCTTCGGAAGC	2169
Qy	201	aAlaLeuGlyLeuPhe-----	206
Db	2170	CTCTCTGGAAAT-TTTCAGGTTCTTTTACTTATTTCTCTCTTAATTTATTTTGTCCCT	2228
Qy	206	-----	206
Db	2229	TTAATATGGTTTTTACTTATGATAAAAGTGTGAGTAATATATTTCGGTCCAATGATGTGG	2288
Qy	206	-----	206
Db	2289	AAAAGATTCCGGCATAAAAAATATCAACCAATAGTAACTTTTTCTTACCTTGTATCTT	2348
Qy	206	-----	206
Db	2349	AAGATCTAAAAATTCATAGCACTATGTACTATGTACTATATGAGAGTGAATATGCGTCAT	2408
Qy	206	-----	206
Db	2409	CGCGAGTTACGTCAACGTGAAAACAAGATACAAAGAGCTGAGACAAACAAATAGGACAA	2468
Qy	206	-----	206
Db	2469	AGCTAGATAACTATTAGAAAAATCACACCGTGGTATGCTCGGTTTAATACATATATACTAC	2528
Qy	206	-----	206
Db	2529	ATGTAACAGTGGAGAAATCATAGAAATAATGTAATTACTTTACTGTAAGTTGGGTATA	2588
Qy	206	-----	206
Db	2589	ATTTAATACATATCTGCGCAAGTTTATTTTGAATTCCTCCCGTATCATTAATCATAGTGTAT	2648
Qy	206	-----	206
Db	2649	GATTCAAAAGCACAGAGAGATATAGTATATATATAGGTAGTATCTTTTATTTTGGTTC	2708
Qy	206	-----	206
Db	2709	CACAAAAATTTTCGGTATATCTAGCCAGAACTGGATTGAAGCATAAAGGTATTAACATAAGTT	2768
Qy	206	-----	206
Db	2769	GGTGAAGTTTGTATGATGATGATGATGATTTTAATAGTGCAAACTGATAAATTTGTTTAGTAG	2828
Qy	207	-----ArgLeuGlyLe	210
Db	2829	CTTTAAAAATAAAATTAAGAGAAGAAATGATGAAAAAGGAATTCGAGGTTAGGTT	2888
Qy	210	uIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAlaAlaThrVa	230
Db	2889	CATAGTGGACTTCTATCGCATGCAACGATAGTAGGATTCATGGGAGGAGCAGCGGT	2948
Qy	230	lValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHisGlyAlaAs	250
Db	2949	GGTGAAGTTCGCAACAGCTTAAGGGTATTTTCGGACTTAACACATTTTCACAGACTCTACCGA	3008
Qy	250	pIleIleSerValMetArgSerValPheThrGlnThrHisGlu-----	264
Db	3009	TGTTATCTCTGTCATGCGTTCCGTTTTTTCCTCCAAACTCAGAGGTCTCTCATTTCCGT	3068
Qy	264	-----	264
Db	3069	TATGTTGCTAATACATCCAAAAATATCATTTTCAATGTCGATGATGTTCTCTCTTAATTAAT	3128
Qy	264	-----	264
Db	3129	ACTATGAATATTTCATAGTATTATTAAACAGTAAACGCTCTAAACATGTTGTTGTTGTT	3188
Qy	265	-----TyrArgTyrGluSerAlaValLeuGlyCysValPheIlePhePh	279
Db	3189	TGTTTTTGTGTGTAGTGGAGATGGGAAAGTGGCGTTCTTGGATGGGTTTTCTATTCTT	3248
Qy	279	eLeuLeuSerThrArgTyrPheSerLys-----	288
Db	3249	TCTTCTCTCCACCATATTTCGT-AAAGATTATAATTTGTTATACCTTTTATATATATCAATTA	3307
Qy	288	-----	288
Db	3308	AATAAGTGAATAATTAACACATTTTATTTCGAATAGCAGCTACATGATCTTTTCGAATA	3367
Qy	288	-----	288
Db	3368	ATTCCATTTACTTTTACTCTTCGTCTTTTTCATCTTTCTCTAGAAAGAAATAAAAAAGTAT	3427
Qy	288	-----	288
Db	3428	ACGATATGGTGGACTGAAAAATGTAAATAACTACACAATCAAAACGTTATTAAATGTGTTT	3487
Qy	289	-----LysArgProArgPhePheTyrValSerAlaMe	299
Db	3488	AATGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3547
Qy	299	tAlaProLeuThrSerValIleLeuGlySerLeuLeuValTyrPheThrHisAlaGlyL	319
Db	3548	GGCTCTTTTGACCTCAGTATGTTCTTGGAAAGTCTCTTGGTTTACTTCTCACTCAGCTGAGAG	3607

QY 319 sHisGlyValGluVal----- 324  
Db 3608 ACATGGTGTCAAGTGGTATTACTCAATTTCTAGTCTAGAAATCTCTTATCTGATGATAA 3667  
QY 324 ----- 324  
Db 3668 ATAATCAAAATAAATAATTTTCAAGTCAAGCAAGCAAGAGATAACGTGTATGATTTCAC 3727  
QY 324 ----- 324  
Db 3728 CGAATCTTAAGAACATTATATATTTTAAAGAAAGAAAGTTAGGTAGAATATTAT 3787  
QY 324 ----- 324  
Db 3788 GGTAAGAAAAGAACAAAGAGCATGAAGTTGAAGTAGTAAGAAATTGTGGAAGAAAG 3847  
QY 324 ----- 324  
Db 3848 TGGGCTCATGGGCTTTTGTGTAGTTTGTTCGAAAGGGAGAGGAATTTATCGAATAAAGGAA 3907  
QY 325 -----IleGlyGluLeuLeuValSerGlyLeuAsnProp 335  
Db 3908 ATTAAGGTATACGATTGTGTCAGATAGGGGACCTGAAAGAAAGGGTTGAATCCAC 3967  
QY 335 roSerLeuThrAsnLeuValPheValSerProTyrMetThrThrAlaValValThrGlyI 355  
Db 3968 TCTCCGGTCTGATCTCACTTTACTTCCCTTACATGTCCACAGCTGTCAAAACTGGCC 4027  
QY 355 leValValGlyIleIleSerLeuAla----- 363  
Db 4028 TCATCACTGGCATCAATGTCTCGCTGTAAAGCCATTATCCCTTTTCTTATTATTAGGCTT 4087  
QY 363 ----- 363  
Db 4088 TGATAGTTATCTTATTATTACCAAGTGAATTAACAAATTAAGGCTTTATACCTA 4147  
QY 363 ----- 363  
Db 4148 AGGCTTAATCACCACACACATATGGGAATATGATTTTCAATGGTGAAGAGTGAAGAAC 4207  
QY 364 -----GluGlyIleAlaValGly 369  
Db 4208 AAAAACTCATCAAAATCAAAATTCATGTGTAAAAATGTAGGAAGAGTAGCAGTGGGG 4267  
QY 370 ArgSerPheAlaMetTyrIleValAsnTyrAsnIleAspGlyAsnIleValMetIleAlaIle 389  
Db 4268 AGAGTTTTGCGATGTTCAGAACTCAACATAGACGGGAACAAAGAGATGATAGCGTTT 4327  
QY 390 GlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThr----- 404  
Db 4328 GGAATGATGAACATCGTTGGTTCCTTCACATCTTGTATTACCTCACAACCGGTACCTACACA 4387  
QY 404 ----- 404  
Db 4388 CTCACACATAGTCACATATGGTTTATTAACCTATGAAGAAAAAGCTTTATATATTATTAC 4447  
QY 405 -----Thr-GlyProPheSerArgSerAlaValAsnTyrAsnAlaGlyCys 419  
Db 4448 GACGTACCATACTATACCAGACCATTTCAAGTCGGCAGTGAACACTCAACCGCGGTTG 4507  
QY 419 sLysThrAlaAlaSerAsnIleIleMetSerLeuAlaValMetLeuThrLeuLeuPheLe 439  
Db 4508 CAAGACCGCAATGTCCCAACATAGTGGGATGGGTTATGTTCCACTCTCTCTCT 4567  
QY 439 uThrProLeuPheHisTyrThrProLeuValValLeuSerAlaIleIleValSerAlaMe 459  
Db 4568 CACACCGCTTTTTCATACACACACTCGTCGTCTCTCTGCGCATCATCATATCCGCAAT 4627  
QY 459 tLeuGlyLeuIleAspTyrGluAlaAlaIleHisLeuPheLeuValAspLysPheAspPh 479  
Db 4628 GCTCGGACTCATTGATATCAAGTCGCCATCCATCTCTGGAAGTTGACAAAGTTCGACTT 4687

QY 479 eValValCysMetSerAlaTyrIleGlyValValPheGlySerValGluIleGlyLeuVa 499  
Db 4688 CCTCGTCTGCATGAGCGCTACGTTGGGCTCGTATTTCGCACTGTAGAGATTGGACTCGT 4747  
QY 499 lIleAla----- 501  
Db 4748 CGTAGCGGTAGTATATATCTTTATAATGCTGAATGAGCCTAGATGCCACTGCTCTTTT 4807  
QY 501 ----- 501  
Db 4808 TTTCCCTTCTATTGCGTTTGGCAATTTTACACTTGCATATATTATAATTTTAAATGTT 4867  
QY 502 -----IleValIleSerValLeuArgValLeuPheIleAl 514  
Db 4868 TTTTITTTTTCATGGTGTAGTGAAGTGGCGATACTATAGCGAGGTGTGCTGTTTGTGTC 4927  
QY 514 aArgProArgThrPheValLeuGlyAsnIleProAsnSerValIleTyrArgAsnValGd 534  
Db 4928 GAGGCCAAAACCTCGGTTGAGGGAAACATACCAACAGCATGATCTATAGGAACACTGA 4987  
QY 534 uHisTyrGlnAsnAlaLysHisValProGlyMetLeuIleLeuGluIleAspAlaProI 554  
Db 4988 GCAGTACCACATCATCAAGAACCGTCTCTGTATTCTCATCTTGGAGATTGATGCTCCCA 5047  
QY 554 eTyrPheAlaAsnAlaSerTyrLeuArgGlu----- 564  
Db 5048 CTACTTTGCTAATGCCAGTTACTTTCGTTGAAAGGTAAACCTAACCTCATTCGTTATTACA 5107  
QY 565 -----ArgIleT 567  
Db 5108 TGCCAAAGATTGAAGAATGTTTTGAATAATGTTTACTTTTCAATAAATCTACAGAATCA 5167  
QY 567 hrArgTyrIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnT 587  
Db 5168 TAAAGTGTGATTGATGAAGAGGAAGAGAGTGTAAACAATCAGAGAGACGCTTACAA 5227  
QY 587 yrValIleIleAspMet----- 592  
Db 5228 ATATTATACTCGATATGTCAGGTAATGCATATGATGATCTTGTGATTCTCTAACACTAGCTAA 5287  
QY 593 -----SerAla 594  
Db 5288 TGTGTTTTTCGTTGGAGGCTTACTTAGTTACTCTCTGCGTGTGTTGTTTAAACAAATCAG 5347  
QY 595 ValGlyAsnIleAspThrSerGlyIleSerMetLeuGluGluValLysLysIleThrGlu 614  
Db 5348 GTTGTATATATCGACAAAGCGGTATTAGCATGATGTTGGTGAATTTAAGAAAGTCATTGAC 5407  
QY 615 ArgArgGluLeu----- 618  
Db 5408 AGGAGAGCGTTTAAAGGTAACGTAGTGCCTGTGTATAATACATTTTATAGGATATGTAACGCGT 5467  
QY 619 -----GlnLeuValLeuValAsnProValIse 627  
Db 5468 TTCTTCTTTGTTAATATTCTTCTTTTAAATCCACAGTTGGTATTGTCAAATCCAAAGG 5527  
QY 627 rGluValMetLysLysLeuAsnLysSerLysPhe--GlnAsnHisLeuGlyLysLysTr 646  
Db 5528 AGAGTCTGTGAAGAAATTAACAGATCCAAATTCATCGGTGATCATTTGGGCAAGAGTGT 5587  
QY 646 pIleTyrLeuThrValGluGluAlaValGlyAlaCysAsnPheAsnLeuArgAlaSerly 666  
Db 5588 GATGTTCTTAAACGTTAGGAGAGCAGTGGAGGCTTGTAGCTACATGCTTCACACGTTTAA 5647  
QY 666 sThrAsnProLysLysAspGluThrGluGlyTyrAsnAsnVal 680  
Db 5648 AACCGAACCG---GCCTCCAAAACAGCGCTTGGAAACACGTA 5687  
RESULT 15  
LOCUS AR477463 2067 bp DNA linear PAT 14-MAY-2004  
DEFINITION Sequence 15 from patent US 6696292.  
ACCESSION AR477463

```

VERSION      AR477463.1  GI:47234921
KEYWORDS     Unknown.
SOURCE       Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 2067)
AUTHORS      Allen,S.M., Falco,S.C. and Thorpe,C.J.
TITLE        Genes encoding sulfate assimilation proteins
JOURNAL      Patent: US 6696292-A 15 24-FEB-2004;
FEATURES     Location/Qualifiers
             1..2067
             source          /organism="unknown"
                             /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.:      2,66e-142      Length:      2067
Score:          1901.00        Matches:     359
Percent Similarity: 76.48%     Conservative: 106
Best Local Similarity: 59.05%  Mismatches:   143
Query Match:     54.71%       Indels:       0
DB:              6            Gaps:         0

US-10-762-049-18 (1-680) x AR477463 (1-2067)

QY      48  ProProGlnProPhePheLysSerLeuLysTyrSerLeuLysGluThrPhePheProAsp 67
      ||| |||
      8  CCACACGAGCACCACTCCACAAACTTAGGCACAGAGTCTCCGAAATCTTCTCCAGAT 67
QY      68  AspProLeuArgGlnPheLysAsnLysProAlaSerLysLysPheMetLeuGlyLeuGln 87
      ||| |||
      68  GACCTCTCCACCGCTTTCAAGAACCACTCGCTTTTAAAGATTCCTCTCGCAGCTTCAG 127
QY      88  PhePhePheProIlePheGluTrpAlaProLysTyrThrPheGlnPheLeuLysAlaAsp 107
      ||| |||
      128  TATCTCTCCCACTTTTCGACTGGGCCCACTACAACTTACCTTACCCTTCTCCGCTCTGAC 187
QY      108  LeuIleAlaGlyIleThrIleAlaSerLeuAlaIleProGlnGlyIleSerTyrAlaLys 127
      ||| |||
      188  CTATCTCTCGGCTTCACCATTCGACGCTCGCATTCCTCAGGGAATCAGTTATGCGAAG 247
QY      128  LeuAlaAsnLeuProIleLeuGlyLeuTyrSerPheIleProProLeuIleTyr 147
      ||| |||
      248  CTTGCCAACTTGGCACCATTCTTGGATTATATTCGAGTTTGTCTCCCAATGATATAC 307
QY      148  AlaMetMetGlySerSerArgAspLeuAlaValGlyThrValAlaValGlySerLeuLeu 167
      ||| |||
      308  TCGCTGCTCGAAGTCTAGACATCTTGGTGTGGACCTGTTTCCATTCGCTCTTTGGTC 367
QY      168  MetGlySerMetLeuSerAsnAlaValAspProAsnGluAspProLysLeuTyrLeuHis 187
      ||| |||
      368  ATGGGATCAATGTTAAGTGATAAAATTTCTTACACTCAAGAACCTATTCTTATCTGGGA 427
QY      188  LeuAlaPheThrAlaThrLeuPheAlaGlyValPheGlnAlaAlaLeuGlyLeuPheArg 207
      ||| |||
      428  TTGGCTTTTACCGCCACTTTCTTTGCTGGGTATTCCAAAGCTTCTCTGGGTATATTAAAG 487
QY      208  LeuGlyLeuIleValAspPheLeuSerHisAlaThrIleIleGlyPheMetGlyGlyAla 227
      ||| |||
      488  CTAGGCTTCGTATTGATTCTTCTGCGAAGCAACGCTGGTGGATTACAGCGGCTGCT 547
QY      228  AlaThrValValCysLeuGlnGlnLeuLysSerIleLeuGlyLeuGluHisPheThrHis 247
      ||| |||
      548  GCCATTATTGTGCTCACTGCGACGAGCTGAAAGGTTTACTTGGAAATAGTGCACTTTTACCAGC 607
QY      248  GlyAlaAspIleIleSerValMetArgSerValPheThrGlnThrHisGluTrpArgTrp 267
      ||| |||
      608  ARGATGCAAAATTAATTCAGTAACGATCTCTGTGTTTCAAGCAAAAGACACAGAGTGGTCATGG 667
QY      268  GluSerAlaValLeuGlyCysValPheIlePhePheLeuLeuSerThrArgTyrPheSer 287
      ||| |||
      668  CAACCACTTCTTTTGGGATTCGGCTTCTCTTGTGCTGACCAACAGGCGACATTAGT 727
```

```

288  LysLysArgProArgPhePheTrpValSerAlaMetAlaProLeuThrSerValIleLeu 307
      ||| |||
      728  TTGAGAAACCAAACTATTCTGGGTTTCAGCAGCTGCCCATTTGACATCAGTTATCTG 787
QY      308  GlySerLeuLeuValTyrPheThrHisAlaGluLysHisGlyValGluValIleGlyGlu 327
      ||| |||
      788  TCACACCATTTTAGTCTTCTCTCGAAATAAGACTCATCAATTTTCAGTTATTTGGGCAC 847
QY      328  LeuLysLysGlyLeuAsnProProSerLeuThrAsnLeuValPheValSerProTyrMet 347
      ||| |||
      848  TTACCAAGAGGAGTTAATCCACCATCAGCAACATGTTTACTTCAATGGTCTCTACTTG 907
QY      348  ThrThrAlaValLysThrGlyIleValValGlyIleIleSerLeuAlaGluGlyIleAla 367
      ||| |||
      908  GGTCTTGTCTATCAAACTGGCATCATCAGAGGATCTTATCTCTACTGAAGGAATTCGA 967
QY      368  ValGlyArgSerPheAlaMetTyrLysAsnTyrAsnIleAspGlyAsnLysGluMetIle 387
      ||| |||
      968  GTAGGGAGAACATTTGCTTCACTTAAGAACTACCAGGTGATGGAAACAAAGAAATGATG 1027
QY      388  AlaIleGlyThrMetAsnValValGlySerPheThrSerCysTyrLeuThrThrGlyPro 407
      ||| |||
      1028  GCCATTGGTCTAATGAACATAGCTGGCTGCTTCTTCTCATGTTATGTTACACGGGATCC 1087
QY      408  PheSerArgSerAlaValAsnTyrAsnAlaGlyCysLysThrAlaAlaSerAsnIleIle 427
      ||| |||
      1088  TTTTCTCGATCGGCTGTTAACTATATATGCTGAGACACAGCAACAGTTTCAATATAATC 1147
QY      428  MetSerLeuAlaValMetLeuThrLeuLeuPheLeuThrProLeuPheHisTyrThrPro 447
      ||| |||
      1148  ATGGCTGACAGCTGTTCTAGTGACATTCGTTTCTCATGCCCTCTTCTACTATACCA 1207
QY      448  LeuValValLeuSerAlaIleIleValSerAlaMetLeuGlyLeuIleAspTyrGluAla 467
      ||| |||
      1208  AATGTTGTCTTAGCGCCATTTATCATCTGCTGTGATTGGTCTAATAGATTATCAATCT 1267
QY      468  AlaIleHisLeuPheLysValAspLysPheAppPheValValCysMetSerAlaTyrIle 487
      ||| |||
      1268  GCATATAAATTGGGAAGGTTGACAACTTGATTCTTCGCCCTGTTTGGCTCTTTT 1327
QY      488  GlyValValPheGlySerValGluIleGlyLeuValIleAlaIleValIleSerValLeu 507
      ||| |||
      1328  GGGTCTCTGTTTCACTTCAGTGGCTTTAGTCTTGGTATAGCGGTTATCATATCAGTCTTC 1387
QY      508  ArgValLeuLeuPheIleAlaArgProArgThrPheValLeuGlyAsnIleProAsnSer 527
      ||| |||
      1388  AAGATCTGCTTCATGCTCACTCGACCAACACTTTTGGTTTGGGGAATATACCAGGAACA 1447
QY      528  ValIleTyrArgAsnValGluHisTyrGlnAsnAlaLysHisValProGlyMetLeuIle 547
      ||| |||
      1448  CAATATTCCCAACACATAAACCAATAACAAAAGCTTTTAAAGAGTTCCTTCTTCTCATTT 1507
QY      548  LeuGluIleAspAlaProIleTyrPheAlaAsnAlaSerTyrLeuArgGluArgIleThr 567
      ||| |||
      1508  TTGGCTGTGTGCTCTCAATCTATTTTGTCTAATCTCACTTATCTTCAAGAAAGGATCTG 1567
QY      568  ArgTyrIleAspGluGluGluArgIleLysAlaThrGlyGluThrSerLeuGlnTyr 587
      ||| |||
      1568  AGATGGTTCGAGAGAGGAGAGGACATATAAAGCTATAATAGGAGTCCATTTGAAGTGC 1627
QY      588  ValIleIleAspMetSerAlaValGlyAsnIleAspThrSerGlyIleSerMetLeuGlu 607
      ||| |||
      1628  ATAATTTTAGATGACAGCTGTCAGCCACAGACACAAAGTGGGCTTGACACTTTTATGT 1687
QY      608  GluValLysLysIleThrGluArgArgGluLeuGlnLeuValLeuValAsnProValSer 627
      ||| |||
      1688  GAACCTTAGAAAGATGCTGGAGAGAGATCAGTTGAGTTGTGCTGGCAATCTCTGTGGA 1747
QY      628  GluValMetLysLysLeuAsnLysSerTyrPheGlnAsnHisLeuGlyLysIleTyrPhe 647
      ||| |||
      1748  AATGTGATGCAAAATTTGCAATAGTCAACACTTTTGGATTCTTTTGGATTAAAGGAGTCT 1807
QY      648  TyrLeuThrValGluGluAlaVal 655
```

Db 1808 TATCTCAGTGGGAGAGCTGTG 1831  
|||||

Search completed: August 31, 2005, 14:27:01  
Job time : 8217 secs